

MỤC LỤC

	Trang
LỜI NÓI ĐẦU	9
PHẦN MỘT: ĐIỀU TRA CHỌN MẪU VÀ SAI SỐ TRONG ĐIỀU TRA THỐNG KÊ	13
1.1. Điều tra chọn mẫu	13
1.1.1. Điều tra chọn mẫu, ưu điểm, hạn chế và điều kiện vận dụng	14
1.1.2. Một số khái niệm và định nghĩa dùng trong điều tra chọn mẫu	18
1.1.3. Xác định cỡ mẫu, phân bổ mẫu và tính sai số chọn mẫu	26
1.2. Sai số trong điều tra thống kê	43
1.2.1. Sai số trong quá trình chuẩn bị điều tra thống kê	44
1.2.2. Sai số trong quá trình tổ chức điều tra	49
1.2.3. Sai số liên quan đến quá trình xử lý thông tin	52
PHẦN HAI: BIỂU HIỆN CÁC MỨC ĐỘ CỦA HIỆN TƯỢNG KINH TẾ - XÃ HỘI	54
2.1. Số tuyệt đối (trong thống kê)	54
2.2. Số tương đối (trong thống kê)	55
2.2.1. Số tương đối động thái	57
2.2.2. Số tương đối so sánh	57
2.2.3. Số tương đối kế hoạch	57
2.2.4. Số tương đối kết cấu	58

2.2.5. Số tương đối cường độ	58
2.3. Số bình quân (trong thống kê)	58
2.3.1. Số bình quân số học	60
2.3.2. Số bình quân điều hoà	61
2.3.3. Số bình quân nhân	62
2.3.4. Mốt	64
2.3.5. Số trung vị	66
2.4. Độ biến thiên của tiêu thức	68
2.4.1. Khoảng biến thiên	68
2.4.2. Độ lệch tuyệt đối bình quân	69
2.4.3. Phương sai	71
2.4.4. Độ lệch chuẩn	72
2.4.5. Hệ số biến thiên	74
2.5. Mức đồng đều của phân phối	75
2.5.1. Đường cong Lorenz	75
2.5.2. Hệ số GINI	77
PHẦN BA: MỘT SỐ PHƯƠNG PHÁP THƯỜNG DÙNG TRONG PHÂN TÍCH THỐNG KÊ	80
3.1. Phương pháp phân tổ thống kê	81
3.1.1. Khái niệm phân tổ thống kê và tiêu thức phân tổ	81
3.1.2. Các loại phân tổ và cách thức tiến hành phân tổ	82
3.2. Phương pháp đồ thị thống kê	85
3.2.1. Biểu đồ hình cột	86

3.2.2. Biểu đồ diện tích	87
3.2.3. Biểu đồ tượng hình	89
3.2.4. Đồ thị đường gấp khúc	90
3.2.5. Biểu đồ hình màng nhện	92
3.3. Phương pháp phân tích dãy số biến động theo thời gian	94
3.3.1. Khái niệm và đặc điểm của dãy số biến động theo thời gian	94
3.3.2. Các chỉ tiêu phân tích dãy số biến động theo thời gian	95
3.3.3. Một số phương pháp biểu hiện xu hướng biến động cơ bản của hiện tượng	101
3.4. Phương pháp phân tích tương quan	110
3.4.1. Liên hệ tương quan và phương pháp phân tích tương quan	110
3.4.2. Phân tích mối liên hệ tương quan giữa các tiêu thức biến đổi theo không gian	111
3.4.3. Phân tích mối liên hệ tương quan giữa hai chỉ tiêu biến động theo thời gian	123
3.5. Phương pháp chỉ số	130
3.5.1. Một số vấn đề chung về phương pháp chỉ số	130
3.5.2. Chỉ số cá thể và chỉ số tổng hợp	133
3.5.3. Chỉ số bình quân	139
3.5.4. Chỉ số liên hoàn và chỉ số định gốc	142
3.5.5. Chỉ số sản phẩm so sánh được và sản phẩm không so sánh được	144
3.5.6. Hệ thống chỉ số	148
3.6. Phương pháp cân đối	152
3.6.1. Bảng cân đối "đơn"	153

3.6.2. Bảng cân đối "kép"	154
---------------------------	-----

PHẦN BỐN: MỘT SỐ CHỈ TIÊU CHỦ YẾU TRONG THỐNG KÊ TÀI KHOẢN QUỐC GIA 156

4.1. Một số khái niệm cơ bản 156

4.1.1. Sản xuất	156
4.1.2. Đơn vị thường trú	157
4.1.3. Đơn vị thể chế	158
4.1.4. Giá cơ bản, giá sản xuất và giá sử dụng	159
4.1.5. Thu nhập sở hữu	162
4.1.6. Chuyển nhượng	162
4.1.7. Biến điểm và biến kỳ	163
4.1.8. Tích sản và tiêu sản	163
4.1.9. Chỉ tiêu cân đối	165

4.2. Một số chỉ tiêu chủ yếu trong thống kê tài khoản quốc gia 165

4.2.1. Giá trị sản xuất	166
4.2.2. Giá trị tăng thêm	168
4.2.3. Tổng sản phẩm trong nước (GDP)	169
4.2.4. Tiêu dùng cuối cùng	171
4.2.5. Tích lũy tài sản	171
4.2.6. Xuất, nhập khẩu hàng hóa và dịch vụ	172
4.2.7. Thu nhập của người lao động từ sản xuất	173
4.2.8. Thuế sản xuất, trợ cấp sản xuất	173
4.2.9. Khấu hao tài sản cố định	174
4.2.10. Thặng dư	175

4.2.11. Tổng thu nhập quốc gia	179
4.2.12. Thu nhập quốc gia thuần	180
4.2.13. Thu nhập quốc gia khả dụng	181
4.2.14. Để dành	182
4.2.15. Thay đổi của cải thuần do thay đổi để dành và chuyển nhượng tài sản	183
4.2.16. Cho vay thuần hay đi vay thuần	184
4.2.17. Bảng tổng kết tài sản	185
4.2.18. Cửa cải thuần	185
4.2.19. Cửa cải thuần đầu kỳ	186
4.2.20. Tích sản phi tài chính cuối kỳ	186
4.2.21. Tích sản tài chính cuối kỳ	187
4.2.22. Tiêu sản cuối kỳ	187
4.2.23. Cửa cải thuần cuối kỳ	188
4.2.24. Thay đổi của cải thuần	188

**PHẦN NĂM: MỘT SỐ CHỈ TIÊU THỐNG KÊ
KINH TẾ - XÃ HỘI TỔNG HỢP**

5.1. Hệ số ICOR	191
5.2. Chỉ số phát triển con người	193
5.3. Chỉ số phát triển giới	196
5.4. Chỉ số biến động về giới	200
5.5. Tốc độ tăng năng suất các nhân tố tổng hợp	204

5.6. Hiệu quả quá trình	209
5.7. Chỉ số thành tựu công nghệ	210
5.8. Chỉ số nghèo tổng hợp	214
TÀI LIỆU THAM KHẢO	217

LỜI NÓI ĐẦU

Để phục vụ cho yêu cầu nghiên cứu, đào tạo cũng như triển khai thực tế về công tác thống kê trong thời kỳ đổi mới, Viện Khoa học Thống kê biên soạn và xuất bản cuốn sách: **"Một số vấn đề phương pháp luận thống kê"**.

Cuốn sách được biên soạn trên cơ sở kế thừa có chọn lọc những vấn đề về phương pháp thống kê truyền thống đã được công bố hoặc đã từng ứng dụng triển thực tế; đồng thời được nghiên cứu cải tiến bổ sung kiến thức thống kê mới trong nước và quốc tế; kết hợp chặt chẽ giữa phương pháp thống kê với phương pháp toán học, giữa nghiên cứu lý luận với tổng kết và ứng dụng thực tiễn; chuẩn hoá khái niệm, định nghĩa, phương pháp tính các chỉ tiêu thống kê, đáp ứng yêu cầu quản lý trong nước và phù hợp với các chuẩn mực thống kê quốc tế, phục vụ việc so sánh trong xu thế đổi mới và hội nhập.

Mặt khác, trong quá trình biên soạn, các tác giả có sử dụng lại một số ví dụ của một số tài liệu đã tính toán để minh chứng cho nội dung và điều kiện áp dụng các phương pháp đã trình bày.

Cuốn sách gồm 5 phần, mỗi phần giới thiệu từng vấn đề về phương pháp luận thống kê riêng biệt, nhưng chúng lại bổ sung cho nhau tạo thành thể thống nhất các phương pháp thống kê.

Phần một với tiêu đề: **"Điều tra chọn mẫu và sai số trong điều tra thống kê"** giới thiệu một cách khái quát có hệ thống những vấn đề cơ bản về lý thuyết chọn mẫu như: Khái niệm, định nghĩa, nội dung điều tra chọn mẫu, ưu điểm, hạn chế và điều kiện vận dụng điều tra chọn mẫu; cách xác định cỡ mẫu, phân bố mẫu và phương pháp tính sai số chọn mẫu,... Trong phần này cũng đề cập tới sai số phi chọn

mẫu xảy ra trong toàn bộ quá trình điều tra thống kê, (Chuẩn bị điều tra, tổ chức thu thập thông tin, tổng hợp số liệu,...). Qua tổng kết thực tiễn điều tra thống kê, cuốn sách đã chỉ rõ sai số phi chọn mẫu ảnh hưởng nhiều đến chất lượng số liệu thống kê và đề xuất những hướng khắc phục nhằm giảm bớt loại sai số này.

Phần hai: **"Biểu hiện các mức độ của hiện tượng kinh tế - xã hội"** đề cập một cách có hệ thống, ngắn gọn, súc tích về phương pháp tính, điều kiện vận dụng các chỉ tiêu phản ánh mức độ và biến động của tiêu thức. Bên cạnh lý thuyết chung, mỗi đại lượng đều có ví dụ minh họa như một tài liệu hướng dẫn nghiệp vụ rõ ràng, thuận tiện cho việc nghiên cứu ứng dụng vào thực tế.

Phần ba đề cập tới **"Một số phương pháp thường dùng trong phân tích thống kê"**. Mỗi phương pháp được trình bày một cách khái quát, tập trung vào những nội dung cơ bản nhất cũng như các hình thức biểu hiện, phương pháp tính và điều kiện vận dụng. Phần này bổ sung một số vấn đề chưa được đề cập trong các tài liệu trước đây hoặc có đề cập nhưng chưa đầy đủ như: Chỉ số sản phẩm so sánh được và sản phẩm không so sánh được; phân tích tương quan dãy số theo thời gian; tự tương quan, đồ thị hình mạng nhện,... vì vậy nội dung các phương pháp phân tích thống kê phong phú và đa dạng hơn, vận dụng vào thực tế thích hợp hơn.

Phần bốn giới thiệu về **"Một số chỉ tiêu chủ yếu trong hệ thống tài khoản quốc gia"**, phần này đề cập một số khái niệm cơ bản dùng trong Hệ thống tài khoản quốc gia SNA làm cơ sở để trình bày ngắn gọn nhưng nêu bật được nội dung, bản chất và mối liên hệ của các chỉ tiêu chủ yếu trong hệ thống tài khoản quốc gia, phản ánh quá trình sản xuất tạo ra thu nhập, phân phối, sử dụng thu nhập cho tiêu dùng, tích lũy, để dành,... Bên cạnh lời văn, cuốn sách đưa ra các công thức mô tả mối liên hệ của các chỉ tiêu này.

Phần cuối của cuốn sách trình bày nội dung phương pháp tính **"Một số chỉ tiêu thống kê kinh tế - xã hội tổng hợp"** thường gặp và

đang là mối quan tâm của người dùng tin. Các chỉ tiêu này được biên soạn độc lập với nhau theo phong cách từ điển. Bên cạnh các chỉ tiêu đã giới thiệu trong cuốn: "Một số thuật ngữ thống kê thông dụng" còn bổ sung các chỉ tiêu thống kê kinh tế - xã hội khác: Tốc độ tăng năng suất các nhân tố tổng hợp, hiệu quả quá trình, Chỉ số thành tựu công nghệ và Chỉ số nghèo tổng hợp. Mỗi chỉ tiêu trình bày đều có ví dụ tính toán khá cụ thể nhằm làm rõ nội dung phương pháp tính, kiểm nghiệm khả năng tính toán và vận dụng của các chỉ tiêu đó.

Với khuôn khổ có hạn, Viện Khoa học Thống kê hy vọng cuốn sách sẽ là tài liệu tham khảo bổ ích, cung cấp những kiến thức cần thiết đáp ứng một phần cho yêu cầu nghiên cứu, đào tạo và vận dụng thực tế trong công tác thống kê. Tuy nhiên, trong quá trình biên soạn và in ấn, cuốn sách không tránh khỏi những hạn chế và sai sót. Viện Khoa học Thống kê mong nhận được góp ý của đông đảo bạn đọc.

Hà Nội, tháng 6 năm 2005

TẬP THỂ TÁC GIẢ

PHẦN MỘT

ĐIỀU TRA CHỌN MẪU VÀ SAI SỐ TRONG ĐIỀU TRA THỐNG KÊ

1.1. ĐIỀU TRA CHỌN MẪU

Quá trình nghiên cứu thống kê gồm các giai đoạn: Thu thập số liệu, xử lý tổng hợp và phân tích, dự báo.

Trong thu thập số liệu thường áp dụng hai hình thức chủ yếu: Báo cáo thống kê định kỳ và điều tra thống kê.

Báo cáo thống kê định kỳ là hình thức thu thập số liệu thống kê được tiến hành thường xuyên, định kỳ theo nội dung, phương pháp cũng như hệ thống biểu mẫu thống nhất, được quy định thành chế độ báo cáo do cơ quan có thẩm quyền quyết định và áp dụng cho nhiều năm.

Điều tra thống kê là hình thức thu thập số liệu được tiến hành theo phương án quy định cụ thể cho từng cuộc điều tra. Trong phương án điều tra quy định rõ mục đích, nội dung, đối tượng, phạm vi, phương pháp và kế hoạch tiến hành điều tra. Điều tra thống kê được áp dụng ngày càng rộng rãi trong điều kiện nền kinh tế thị trường có nhiều thành phần kinh tế.

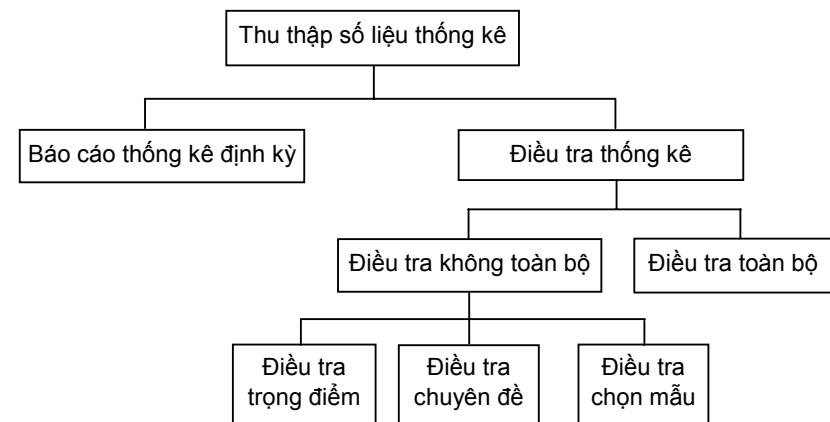
Điều tra thống kê được phân thành điều tra toàn bộ và điều tra không toàn bộ. Điều tra toàn bộ nhằm tiến hành thu thập số liệu ở tất cả các đơn vị của tổng thể. Trong khi đó điều tra không toàn bộ chỉ tiến hành thu thập số liệu của một bộ phận các đơn vị trong tổng thể. Trong điều tra không toàn bộ còn chia ra điều tra trọng điểm, điều tra chuyên đề và điều tra chọn mẫu.

Điều tra trọng điểm và điều tra chuyên đề khác với điều tra chọn

mẫu ở chỗ kết quả của nó không dùng để suy rộng cho tổng thể chung. Kết quả của điều tra chọn mẫu được dùng để mô tả đặc điểm của tổng thể chung.

Các hình thức thu thập số liệu thống kê trên đây có thể khái quát qua sơ đồ sau:

Sơ đồ 1.1. Các hình thức và phương pháp thu thập số liệu thống kê



Dưới đây đi sâu nghiên cứu "Điều tra chọn mẫu".

1.1.1. Điều tra chọn mẫu, ưu điểm, hạn chế và điều kiện vận dụng

1.1.1.1. Khái niệm điều tra chọn mẫu

Điều tra chọn mẫu (ĐTCM) là loại điều tra không toàn bộ, trong đó người ta chọn một cách ngẫu nhiên một số đủ lớn đơn vị đại diện trong toàn bộ các đơn vị của tổng thể chung để điều tra rồi dùng kết quả thu thập được tính toán, suy rộng thành các đặc điểm của toàn bộ tổng thể chung. Ví dụ, để có năng suất và sản lượng lúa của một địa

bàn điều tra nào đó (huyện A chẳng hạn) người ta chỉ tiến hành thu thập số liệu về năng suất và sản lượng lúa thu trên diện tích của một số hộ gia đình được chọn vào mẫu của huyện để điều tra thực tế, sau đó dùng kết quả thu được tính toán và suy rộng cho năng suất và sản lượng lúa của toàn huyện A.

ĐTCM được ứng dụng rất rộng rãi trong thống kê kinh tế - xã hội như: Điều tra năng suất, sản lượng lúa; Điều tra lao động - việc làm; Điều tra thu nhập, chi tiêu của hộ gia đình; Điều tra biến động thường xuyên dân số; Điều tra chất lượng sản phẩm công nghiệp.

Ngoài ra, trong tự nhiên, trong đời sống sinh hoạt của con người, trong y học, v.v... chúng ta cũng đã gặp rất nhiều ví dụ thực tế đã áp dụng ĐTCM; chẳng hạn: Khi đo lượng nước mưa của một khu vực nào đó người ta chỉ chọn ra một số điểm trong khu vực và đặt các ống nghiệm (các mẫu) để đo lượng nước mưa qua các trận mưa trong từng tháng và cả năm, sau đó dựa vào kết quả nước mưa đo được từ mẫu là các ống nghiệm để tính toán suy rộng về lượng nước trung bình các tháng và cả năm cho cả khu vực; khi nghiên cứu ảnh hưởng của hút thuốc lá đối với sức khỏe con người, người ta chọn ra một số lượng cần thiết người hút thuốc lá để kiểm tra sức khỏe và dùng kết quả kiểm tra từ một số người đó để kết luận về ảnh hưởng của hút thuốc lá tới sức khỏe cộng đồng, v.v...

1.1.1.2. Ưu điểm của điều tra chọn mẫu

Do chỉ tiến hành điều tra trên một bộ phận đơn vị mẫu trong tổng thể chung nên ĐTCM có những ưu điểm cơ bản sau:

- Tiến hành điều tra nhanh gọn, bảo đảm tính kịp thời của số liệu thống kê.
- Tiết kiệm nhân lực và kinh phí trong quá trình điều tra.
- Cho phép thu thập được nhiều chỉ tiêu thống kê, đặc biệt đối với các chỉ tiêu có nội dung phức tạp, không có điều kiện điều tra ở diện rộng. Nhờ đó kết quả điều tra thu được sẽ phản ánh được nhiều mặt,

cho phép nghiên cứu các mối quan hệ cần thiết của hiện tượng nghiên cứu.

- Làm giảm sai số phi chọn mẫu (sai số do cân, đong, đo, đếm, khai báo, ghi chép, v.v...). Trong thực tế công tác thống kê sai số phi chọn mẫu luôn luôn tồn tại và ảnh hưởng không nhỏ đến chất lượng số liệu thống kê, nhất là các chỉ tiêu có nội dung phức tạp, việc tiếp cận để thu thập số liệu khó khăn, tốn nhiều thời gian trong quá trình phỏng vấn, ghi chép và đặc biệt hơn là đối với các chỉ tiêu điều tra không có sẵn thông tin mà đòi hỏi phải hỏi tường để nhớ lại. Đối với những loại thông tin như trên, chỉ có tiến hành điều tra mẫu mới có điều kiện tuyển chọn điều tra viên tốt hơn; hướng dẫn nghiệp vụ kỹ hơn, thời gian dành cho một đơn vị điều tra nhiều hơn, tạo điều kiện cho các đối tượng cung cấp thông tin trả lời chính xác hơn, tức là làm cho sai số phi chọn mẫu ít hơn.

- Cho phép nghiên cứu các hiện tượng kinh tế - xã hội, môi trường,... không thể tiến hành theo phương pháp điều tra toàn bộ: Ví dụ như nghiên cứu trữ lượng khoáng sản, thủy sản,...

1.1.1.3. Hạn chế của điều tra chọn mẫu

- Do ĐTCM chỉ tiến hành thu thập số liệu trên một số đơn vị, sau đó dùng kết quả để suy rộng cho toàn bộ tổng thể chung nên kết quả điều tra chọn mẫu luôn tồn tại cái gọi là "Sai số chọn mẫu" - Sai số do tính đại diện. Sai số chọn mẫu phụ thuộc vào độ đồng đều của chỉ tiêu nghiên cứu, vào cỡ mẫu và phương pháp tổ chức chọn mẫu. Có thể làm giảm sai số chọn mẫu bằng cách tăng cỡ mẫu ở phạm vi cho phép và lựa chọn phương pháp tổ chức chọn mẫu thích hợp nhất.

- Kết quả ĐTCM không thể tiến hành phân nhỏ theo mọi phạm vi và tiêu thức nghiên cứu như điều tra toàn bộ, mà chỉ thực hiện được ở mức độ nhất định tùy thuộc vào cỡ mẫu, phương pháp tổ chức chọn mẫu và độ đồng đều giữa các đơn vị theo các chỉ tiêu được điều tra.

1.1.1.4. Điều kiện vận dụng của điều tra chọn mẫu

Điều tra chọn mẫu thường được vận dụng trong các trường hợp sau:

- Thay thế cho điều tra toàn bộ trong những trường hợp quy mô điều tra lớn, nội dung điều tra cần thu thập nhiều chỉ tiêu, thực tế ta không đủ kinh phí và nhân lực để tiến hành điều tra toàn bộ, hơn nữa nếu điều tra toàn bộ sẽ mất quá nhiều thời gian, không đảm bảo tính kịp thời của số liệu thống kê như điều tra thu nhập, chi tiêu hộ gia đình, điều tra năng suất, sản lượng lúa, điều tra vốn đầu tư của các đơn vị ngoài quốc doanh...; hoặc không tiến hành được điều tra toàn bộ vì không thể xác định được tổng thể chung như điều tra đánh giá mức độ ô nhiễm môi trường nước của một số sông, hồ nào đó (tổng thể chung phải là toàn bộ lượng nước có trong các sông, hồ được xác định là đã bị ô nhiễm),...

- Quá trình điều tra gắn liền với việc phá hủy sản phẩm như điều tra đánh giá chất lượng thịt hộp, cá hộp, đánh giá chất lượng đạn dược, y tá lấy máu của bệnh nhân để xét nghiệm, v.v... Các trường hợp trên đây nếu điều tra toàn bộ thì sau khi điều tra toàn bộ sản phẩm sản xuất ra hoặc lượng máu có trong cơ thể của bệnh nhân sẽ bị phá hủy hoàn toàn. Đây là điều không bao giờ cho phép thực hiện trong thực tế.

- Để thu thập những thông tin tiên nghiệm trong những trường hợp cần thiết nhằm phục vụ cho yêu cầu của điều tra toàn bộ. Ví dụ, để thăm dò mức độ tín nhiệm của các ứng cử viên vào một chức vị nào đó thì chỉ có thể ĐTCM ở một lượng cử tri nhất định và phải được tiến hành trước khi bầu cử chính thức thì mới có ý nghĩa (Bỏ phiếu bầu cử chính thức chính là điều tra toàn bộ).

- Thu thập số liệu để kiểm tra, đánh giá và chỉnh lý số liệu của điều tra toàn bộ. Trong thực tế có những cuộc điều tra toàn bộ có quy mô lớn hoặc điều tra rất phức tạp như Tổng Điều tra Dân số và Nhà ở, Tổng Điều tra Nông thôn, Nông nghiệp và Thủy sản,... thì sai số do khai báo, thu thập thông tin thường xuyên tồn tại và ảnh hưởng đáng

kể đến chất lượng số liệu. Vì vậy cần có ĐTCM với quy mô nhỏ hơn để xác định mức độ sai số này, trên cơ sở đó tiến hành đánh giá độ tin cậy của số liệu và nếu ở mức độ cần thiết có thể phải chỉnh lý lại số liệu thu được từ điều tra toàn bộ.

1.1.2. Một số khái niệm và định nghĩa dùng trong điều tra chọn mẫu

1.1.2.1. Tổng thể chung và tổng thể mẫu⁽¹⁾

a. Các tham số của tổng thể chung

Tổng thể chung là toàn bộ các đơn vị thuộc đối tượng điều tra của một cuộc ĐTCM.

Gọi U_i ($i = 1, 2, \dots, N$) là các đơn vị thuộc đối tượng điều tra với X_i là trị số tiêu thức nghiên cứu của từng đơn vị tổng thể, thì toàn bộ các U_i là tổng thể chung. Và khi đó sẽ có công thức tính các tham số:

- Giá trị của tổng thể chung:

$$X = X_1 + X_2 + \dots + X_N = \sum_{i=1}^N X_i \quad ; \quad (1.1.1)$$

- Đại lượng bình quân của tổng thể chung:

$$\bar{X} = \frac{X}{N} = \frac{1}{N} \sum_{i=1}^N X_i \quad ; \quad (1.1.2)$$

- Phương sai của tổng thể chung:

$$S^2 = \frac{1}{N} \sum_{i=1}^N (X_i - \bar{X})^2 \quad ; \quad (1.1.3)$$

b. Các tham số của tổng thể mẫu

⁽¹⁾ Ở đây chỉ đề cập trường hợp điều tra nghiên cứu chỉ tiêu bình quân làm ví dụ.

Tổng thể mẫu là bộ phận của tổng thể chung gồm những đơn vị được lựa chọn để trực tiếp thu thập thông tin trong một cuộc điều tra chọn mẫu.

Gọi u_i ($i = 1, 2, \dots, n$) là các đơn vị thuộc đối tượng điều tra được chọn vào mẫu, với x_i là trị số tiêu thức nghiên cứu từng đơn vị mẫu, thì toàn bộ u_i là tổng thể mẫu và n là số đơn vị tổng thể mẫu. Tổng thể mẫu có các tham số tính theo phạm vi tổng thể mẫu như sau:

- Giá trị của tổng thể mẫu:

$$x = x_1 + x_2 + \dots + x_n = \sum_{i=1}^n x_i \quad ; \quad (1.1.4)$$

- Đại lượng bình quân mẫu:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{x}{n} \quad ; \quad (1.1.5)$$

- Phương sai mẫu điều chỉnh (gọi tắt là phương sai mẫu):

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad ; \quad (1.1.6)$$

1.1.2.2. Ước lượng

Nội dung cơ bản của phương pháp điều tra chọn mẫu là dựa vào sự hiểu biết về tham số θ' nào đó của tổng thể mẫu đã điều tra để suy luận thành tham số θ của tổng thể chung. Việc suy luận đó gọi là ước lượng.

a. Tiêu chuẩn của ước lượng

Có ước lượng chệch và ước lượng không chệch. Tham số θ' của tổng thể mẫu được gọi là ước lượng không chệch của tham số θ của tổng thể chung nếu $M(\theta') = \theta$ (kỳ vọng toán của θ' bằng θ). Nếu ước lượng không thoả mãn điều kiện trên được gọi là ước lượng chệch.

Thống kê toán đã chứng minh và rút ra một số kết luận sau:

+ Vì số bình quân mẫu $\bar{x}$ là ước lượng không chệch, hiệu quả và vững của số bình quân tổng thể chung $\bar{x}$, do đó nếu chưa biết $\bar{x}$ có thể dùng $\bar{x}$ để ước lượng.

+ Vì phương sai điều chỉnh mẫu s^2 là ước lượng không chệch, hiệu quả và vững của phương sai chung S^2 , do đó nếu chưa biết phương sai S^2 có thể dùng s^2 để ước lượng.

b. Các phương pháp ước lượng

Có 2 phương pháp sử dụng θ' để ước lượng θ : Phương pháp ước lượng điểm và phương pháp ước lượng bằng khoảng tin cậy.

- Phương pháp ước lượng điểm là dùng một tham số của mẫu để suy luận cho tham số θ chưa biết của tổng thể chung vì bản thân θ là một số xác định.

- Phương pháp ước lượng bằng khoảng tin cậy là từ một tham số θ' của tổng thể mẫu xây dựng một khoảng giá trị (θ'_1, θ'_2) sao cho với một xác suất cho trước, tham số θ sẽ rơi vào khoảng (θ'_1, θ'_2) đó, hay nói cách khác là khoảng (θ'_1, θ'_2) sẽ chứa đựng giá trị θ với một xác suất cho trước. Khoảng (θ'_1, θ'_2) của tham số tổng thể mẫu được gọi là khoảng tin cậy của tham số tổng thể chung θ nếu với xác suất bằng $(1 - \alpha)$ cho trước thoả mãn điều kiện:

$$P(\theta'_2 < \theta < \theta'_1) = 1 - \alpha ;$$

$(1 - \alpha)$ được gọi là xác suất tin cậy của ước lượng, $I = \theta'_2 - \theta'_1$ được gọi là khoảng tin cậy.

1.1.2.3. Sai số chọn mẫu và phạm vi sai số chọn mẫu

a. Sai số chọn mẫu

Sai số chọn mẫu (SSCM) là sự khác nhau giữa giá trị ước lượng của mẫu và giá trị của tổng thể chung. Sai số chọn mẫu còn gọi là sai số do tính đại diện. Sai số này chỉ xảy ra trong điều tra chọn mẫu do chỉ điều tra một số ít đơn vị mà kết quả lại suy cho cả tổng thể. Sai số

chọn mẫu có hai loại:

- Sai số có hệ thống: Sai số xảy ra khi áp dụng phương pháp chọn có hệ thống, làm cho kết quả điều tra luôn bị lệch so với số thực tế về một hướng.

- Sai số ngẫu nhiên: Sai số chỉ xuất hiện trong trường hợp các đơn vị của tổng thể được chọn theo nguyên tắc ngẫu nhiên, không phụ thuộc vào ý định của người điều tra.

b. Phạm vi sai số chọn mẫu

Phạm vi SSCM (ký hiệu là Δ_x) bằng tích của hệ số tin cậy (t) và SSCM (μ_x)

$$\Delta_x = t \cdot \mu_x \quad ; \quad (1.1.7)$$

Trong đó: Hệ số tin cậy (tương ứng với độ tin cậy ϕ_t) là xác suất để giá trị thực tế của chỉ tiêu nghiên cứu ($\bar{X}$) còn nằm trong khoảng tin cậy ($\bar{x} - t \cdot \mu_x$ đến $\bar{x} + t \cdot \mu_x$).

Theo chứng minh của toán học thì t tương ứng với hàm xác suất (ϕ_t) đã được Li -a-pu-nốp tính sẵn và lập thành bảng. Ý nghĩa của hàm xác suất này được biểu hiện như sau:

$$P\left[|\bar{x} - \bar{X}| \leq \Delta_x\right] = \phi_{(t)} = 1 - \alpha$$

Sau đây là một vài trị số tiêu biểu:

$t = 1$ thì $\phi_t = 0,6827$; $t = 2$ thì $\phi_t = 0,9545$; $t = 3$ thì $\phi_t = 0,9973$

Như vậy, có thể ước lượng tham số của tổng thể chung bằng khoảng tin cậy với công thức như sau:

$$\bar{X} = \bar{x} \pm \Delta_x \Rightarrow \bar{x} - \Delta_x \leq \bar{X} \leq \bar{x} + \Delta_x ; \quad (1.1.8)$$

c. Ý nghĩa của việc tính toán sai số chọn mẫu

- Sai số chọn mẫu dùng để ước lượng chỉ tiêu nghiên cứu theo khoảng tin cậy, điều này thể hiện qua công thức 1.1.8.

- Sai số chọn mẫu dùng để đánh giá tính đại diện của chỉ tiêu

nghiên cứu qua tính toán tỷ lệ SSCM (H) như sau:

$$H = \frac{\mu}{\bar{x}} \times 100 \quad ; \quad (1.1.9)$$

H càng nhỏ thì chỉ tiêu có tính đại diện càng cao và ngược lại.

- Là cơ sở để xác định cỡ mẫu cho các cuộc điều tra được tiến hành về sau.

1.1.2.4. Đơn vị chọn mẫu và dàn chọn mẫu

a. Đơn vị chọn mẫu

Đơn vị chọn mẫu là các đơn vị cơ bản hoặc nhóm đơn vị cơ bản được xác định rõ ràng, tương đối đồng đều và có thể quan sát được, thích hợp cho mục đích chọn mẫu. Ví dụ: Doanh nghiệp, hộ gia đình, đơn vị diện tích gieo trồng, xã, phường, xóm, bản...

Nếu chọn mẫu một cấp thì có một loại đơn vị chọn mẫu, còn nếu chọn mẫu nhiều cấp thì sẽ có nhiều loại đơn vị chọn mẫu. Tức là lược đồ chọn mẫu theo bao nhiêu cấp thì có bấy nhiêu loại đơn vị chọn mẫu.

b. Dàn chọn mẫu

Dàn chọn mẫu có thể là danh sách các đơn vị chọn mẫu với những đặc điểm nhận dạng của chúng hoặc là bản đồ chỉ ra ranh giới của các đơn vị được dùng làm căn cứ để tiến hành chọn mẫu. Khi tổ chức điều tra thống kê.

Trong tổng thể nghiên cứu, tùy thuộc vào lược đồ chọn mẫu mà sẽ có các loại dàn chọn mẫu khác nhau. Nếu điều tra mẫu một cấp (giả định điều tra các hộ trên địa bàn huyện) thì dàn chọn mẫu là danh sách các hộ gia đình của tất cả các xã trong huyện. Còn nếu điều tra mẫu hai cấp, cấp I là xã và cấp II là hộ gia đình thì có hai loại dàn chọn mẫu: Dàn chọn mẫu cấp I là danh sách tất cả các xã trong huyện, còn dàn chọn mẫu cấp II là danh sách các hộ gia đình của những xã được chọn ở mẫu cấp I.

1.1.2.5. Chọn mẫu ngẫu nhiên, chọn mẫu hệ thống và chọn theo phương pháp phân tích chuyên gia

- Chọn mẫu ngẫu nhiên là chọn các đơn vị từ tổng thể vào mẫu hoàn toàn hù hoạ. Cách đơn giản nhất của chọn mẫu ngẫu nhiên là rút thăm hoặc sử dụng bảng số ngẫu nhiên.

- Chọn mẫu hệ thống là chọn các đơn vị từ tổng thể vào mẫu theo một khoảng cách cố định sau khi đã chọn ngẫu nhiên một nhóm nào đó trên cơ sở các đơn vị điều tra được sắp xếp thứ tự theo một tiêu thức nhất định.

Vi dụ: Trường đại học "X" có 2000 sinh viên ($N = 2000$). Cần chọn 100 sinh viên ($n = 100$) để điều tra mức sống của họ. Nếu chọn hệ thống sẽ tiến hành như sau:

+ Lập danh sách 2000 sinh viên của trường theo thứ tự nào đó, chẳng hạn theo vần A, B, C... của tên gọi.

+ Chia tổng số sinh viên của trường thành 100 nhóm đều nhau và sẽ có số sinh viên mỗi nhóm là 20 sinh viên: ($K = N: n = 2000 : 100$).

+ Chọn ngẫu nhiên một sinh viên ở nhóm thứ nhất, chẳng hạn rơi vào sinh viên có số thứ tự 15.

+ Mỗi nhóm khác còn lại sẽ chọn 1 sinh viên có số thứ tự: nhóm 2: $(15+K)$, nhóm 3: $(15+2K)$,..., nhóm 100: $(15+99K)$.

Kết quả chọn được 100 sinh viên như vậy được gọi là chọn hệ thống.

- Chọn mẫu theo phương pháp phân tích chuyên gia là chọn mẫu trên cơ sở phân tích xem xét chủ quan của người điều tra. Cách chọn này thường áp dụng cho tổng thể có ít đơn vị mẫu hoặc trị số của chỉ tiêu nghiên cứu giữa các đơn vị mẫu chênh lệch nhau nhiều.

1.1.2.6. Các phương pháp tổ chức chọn mẫu

Có nhiều phương pháp, tổ chức chọn mẫu khác nhau. Mỗi phương pháp có những ưu, nhược điểm riêng và được áp dụng trong những điều kiện nhất định. Tuy nhiên gọi là phương pháp này hay phương pháp kia là đứng trên những giác độ khác nhau và cũng chỉ có ý nghĩa tương đối.

- Xét theo cấp chọn mẫu có phương pháp tổ chức chọn mẫu một cấp và tổ chức chọn mẫu hai cấp hay nhiều cấp:

+ Chọn mẫu một cấp là từ một loại danh sách của tất cả các đơn vị thuộc tổng thể chung, tiến hành chọn mẫu một lần trực tiếp đến các đơn vị điều tra không qua một phân đoạn nào khác.

Chọn mẫu một cấp chỉ có một loại đơn vị chọn mẫu và một dàn chọn mẫu. Đối với mẫu một cấp có thể dùng cách chọn ngẫu nhiên, nhưng cũng có thể dùng cách chọn hệ thống hoặc chọn theo phương pháp chuyên gia. Tuy nhiên, trong thực tế nếu là điều tra mẫu một cấp thì phổ biến là dùng cách chọn ngẫu nhiên và thường được gọi tắt là "chọn mẫu ngẫu nhiên đơn giản". Chọn mẫu ngẫu nhiên đơn giản đảm bảo số mẫu được rải trên toàn địa bàn điều tra nên SSCM sẽ nhỏ. Song khó khăn là việc lập danh sách các đơn vị (dàn chọn mẫu) để tiến hành chọn mẫu khá lớn, tốn nhiều thời gian và công sức. Hơn nữa khi tổ chức điều tra phải thực hiện ở địa bàn rất rộng.

+ Chọn mẫu nhiều cấp là tiến hành điều tra theo nhiều công đoạn, trong đó mỗi công đoạn là một cấp chọn mẫu. Có bao nhiêu cấp điều tra thì có bấy nhiêu loại đơn vị chọn mẫu cũng như có bấy nhiêu loại dàn chọn mẫu.

Phương pháp tổ chức chọn mẫu nhiều cấp thuận tiện cho việc lập dàn chọn mẫu và tổ chức điều tra: Ở cấp sau chỉ phải lập dàn chọn mẫu cho cấp đó trong phạm vi mẫu cấp trước được chọn, phạm vi điều tra được thu hẹp sau mỗi cấp điều tra. Tuy nhiên, với phương pháp tổ chức chọn mẫu nhiều cấp số liệu thu thập được thường có độ tin cậy thấp hơn so với chọn mẫu ngẫu nhiên đơn giản.

- Nếu trước khi chọn mẫu, tiến hành phân chia tổng thể thành

những tổ khác nhau theo một hay một số tiêu thức nào đó liên quan đến tiêu thức điều tra, sau đó phân bổ cỡ mẫu cho từng tổ và trong mỗi tổ lập một danh sách riêng và chọn đủ số mẫu phân bổ cho tổ đó. Cách chọn như vậy gọi là chọn mẫu phân tổ.

Với phương pháp chọn mẫu phân tổ, nếu việc phân tổ được tiến hành khoa học thì tổng thể mẫu sẽ có kết cấu gần tổng thể chung, do đó SSCM sẽ giảm đi, tính chất đại diện của tổng thể mẫu được nâng cao.

Tuy nhiên, chọn mẫu phân tổ cũng khó khăn trong việc lập dàn chọn mẫu như chọn mẫu ngẫu nhiên đơn giản. Hơn nữa tổ chức điều tra phải tiến hành trên địa bàn rộng, thậm chí còn phức tạp hơn cả chọn mẫu ngẫu nhiên đơn giản.

- Nếu điều tra chia thành nhiều cấp, các cấp tiến hành trước thì chọn từng đơn vị mẫu, nhưng ở cấp cuối cùng không chọn ra từng đơn vị, mà chọn cả nhóm các đơn vị để điều tra. Cách chọn như vậy gọi là chọn mẫu chùm (hay chọn mẫu cả khối).

Nếu cùng cỡ mẫu như nhau, chọn mẫu chùm so với các phương pháp tổ chức chọn mẫu nêu trên sẽ thuận tiện nhất cho việc lập dàn chọn mẫu và tổ chức điều tra. Tuy nhiên, độ tin cậy của số liệu thu thập được sẽ thấp hơn; tức là có SSCM lớn nhất.

1.1.3. Xác định cỡ mẫu, phân bổ mẫu và tính sai số chọn mẫu

1.1.3.1. Xác định cỡ mẫu (số đơn vị mẫu)

Xác định cỡ mẫu (số đơn vị mẫu) chính là xác định số lượng đơn vị điều tra trong tổng thể mẫu để tiến hành thu thập số liệu. Yêu cầu của cỡ mẫu là vừa đủ để vừa đảm bảo độ tin cậy cần thiết của số liệu điều tra vừa đảm bảo phù hợp với điều kiện về nhân lực và kinh phí và có thể thực hiện được, tức là có tính khả thi.

Dưới đây sẽ trình bày cách xác định cỡ mẫu đơn thuần theo lý thuyết và việc xác định cỡ mẫu trong thực tế các cuộc điều tra thống kê ở Việt Nam.

a. *Xác định cỡ mẫu theo các công thức lý thuyết.* Một tổng thể khi tiến hành điều tra không chia thành các tổng thể nhỏ (các tổ) thì chỉ có một cách xác định cỡ mẫu trên cơ sở thông tin về quy mô và phương sai của tổng thể chung. Đối với một tổng thể khi điều tra có chia thành các tổng thể nhỏ có hai cách xác định cỡ mẫu: Cách thứ nhất xác định cỡ mẫu như trường hợp không phân tổ, sau đó phân bổ số mẫu chung cho các tổ theo nguyên tắc phân bổ mẫu. Cách thứ hai xác định cỡ mẫu trên cơ sở quy mô và phương sai của từng tổ.

Sau đây sẽ giới thiệu công thức xác định cỡ mẫu theo hai cách nói trên nhưng chỉ cho trường hợp tổ chức chọn mẫu ngẫu nhiên đơn giản hoặc có phân tổ và được áp dụng cho nghiên cứu chỉ tiêu bình quân với cách chọn không lập làm ví dụ.

+ Cách thứ nhất xác định cỡ mẫu trên cơ sở các thông tin về quy mô và phương sai của tổng thể chung:

$$n = \frac{N \cdot t^2 \cdot S^2}{N \cdot \Delta_x^2 + t^2 \cdot S^2} \quad ; \quad (1.1.10)$$

Trong đó:

N - Số đơn vị tổng thể chung;

n - Số đơn vị mẫu;

t - Hệ số tin cậy;

Δ_x - Phạm vi sai số chọn mẫu;

S^2 - Phương sai của tổng thể chung.

+ Cách thứ hai xác định cỡ mẫu trên cơ sở các thông tin về quy mô và phương sai của các tổ t:

$$n = \frac{\sum_{t=1}^K w_t S_t^2}{\frac{\Delta_x^2}{t_\alpha^2} + \frac{1}{N} \sum_{t=1}^K w_t S_t^2} \quad ; \quad (1.1.11)$$

Trong đó:

N - Số đơn vị tổng thể chung;

n - Số đơn vị mẫu;

t_α - Hệ số tin cậy;

Δ_x - Phạm vi sai số chọn mẫu;

w_t - Tỷ trọng số đơn vị của tổ t trong tổng thể chung;

K - Số lượng tổ ($t = 1, 2, \dots, K$);

S_t^2 - Phương sai tổng thể chung của tổ t.

Từ các công thức trên, để xác định cỡ mẫu trong quá trình chuẩn bị phương án điều tra phải có được những thông tin sau:

- N: Số đơn vị tổng thể. Chỉ tiêu này có đầy đủ ở phần lớn các cuộc điều tra thống kê;

- w_t : Tỷ trọng số đơn vị của tổ t trong tổng thể. Đại lượng này xác định được trên cơ sở so sánh số đơn vị từng tổ (N_t) với số đơn vị toàn bộ tổng thể (N);

- t_α , Δ_x : Hệ số tin cậy và phạm vi sai số chọn mẫu là những thông tin của chỉ tiêu điều tra và được ấn định từ trước do yêu cầu thuộc chủ quan của những người quản lý và tổ chức điều tra;

- S_t^2 : Phương sai của từng tổ t. Số liệu để tính các phương sai trên, cần có trước khi điều tra, song thực tế lại không có, do vậy thường phải dùng số liệu điều tra toàn bộ của các cuộc điều tra trước (nếu có). Trường hợp không có số liệu của các cuộc điều tra trước thì phải tiến hành điều tra mẫu nhỏ. Tuy nhiên, việc điều tra mẫu nhỏ cũng khá phức tạp, mất nhiều thời gian, nhiều khi còn ảnh hưởng đến tiến độ thực hiện của cuộc điều tra chính.

Một khó khăn nữa là trong một cuộc ĐTCM thường tiến hành thu thập thông tin về nhiều chỉ tiêu. Các chỉ tiêu khác nhau sẽ có quy luật phân phối và độ biến thiên khác nhau, tức là có phương sai khác nhau.

Và do vậy, mỗi chỉ tiêu tính ra sẽ có một cỡ mẫu riêng (mặc dù yêu cầu về độ tin cậy (ϕ_t) của các chỉ tiêu điều tra như nhau). Nói cách khác, có bao nhiêu chỉ tiêu điều tra thì phải tính bấy nhiêu cỡ mẫu, sau đó sẽ chọn ra cỡ mẫu lớn nhất dùng chung cho điều tra tất cả các chỉ tiêu. Với nhiều cỡ mẫu đòi hỏi phải tính nhiều phương sai nên công việc tính toán càng trở nên phức tạp, tốn nhiều công sức, khó thực hiện.

Vì những đặc điểm trên đây, trong thực tế điều tra chọn mẫu ở nước ta còn ít khi áp dụng một cách trực tiếp các công thức trên để xác định cỡ mẫu.

Ngành Thống kê trong những năm gần đây đã có một số cuộc điều tra chọn mẫu mà các chuyên gia chọn mẫu đã dựa vào thông tin của các cuộc điều tra có liên quan trước đó để xác định cỡ mẫu theo công thức lý thuyết. Song kết quả thu được còn khiêm tốn.

b. Xác định cỡ mẫu theo kinh nghiệm điều tra thực tế. Trong thực tế nhiều khi các chuyên gia thống kê thường căn cứ vào cỡ mẫu của các cuộc điều tra có điều kiện và quy mô tương tự đã thực hiện thành công trước đó ở trong nước hoặc trên thế giới để xác định cỡ mẫu cho cuộc điều tra sau. Có nhiều cách xác định cỡ mẫu nhưng phổ biến nhất vẫn dựa vào tỷ lệ mẫu chung đã được điều tra và bổ sung thêm một tỷ lệ mẫu dự phòng nào đó.

Cách làm này đơn giản, nhanh chóng và dễ thực hiện, tức là có tính khả thi cao. Tuy nhiên làm như vậy chủ yếu vẫn là theo chủ nghĩa kinh nghiệm và gần như chưa tính đến mức độ biến động của các chỉ tiêu nghiên cứu.

c. Xác định cỡ mẫu cũng dựa theo cỡ mẫu của cuộc điều tra nào đó (có điều kiện, quy mô tương tự và đã được tiến hành thành công), *nhưng có điều chỉnh* (tăng lên hoặc giảm đi) trên cơ sở phân tích tỷ lệ SSCM của một số chỉ tiêu chủ yếu. Quá trình này được tiến hành theo hai hướng:

Trước hết liệt kê những chỉ tiêu chủ yếu cùng được tổ chức thu thập số liệu trong cả 2 cuộc điều tra (cuộc điều tra trước đó đã hoàn chỉnh và cuộc điều tra lần này đang chuẩn bị); trong đó chọn ra một chỉ tiêu trong cuộc điều tra lần trước có tỷ lệ SSCM lớn nhất (từ đây chỉ tiêu được chọn gọi là chỉ tiêu nghiên cứu).

Tiếp theo, tiến hành xem xét tỷ lệ SSCM của chỉ tiêu nghiên cứu tính được của cuộc điều tra lần trước và xử lý như sau:

- Nếu tỷ lệ SSCM đó lớn hơn mức độ cho phép thì phải điều chỉnh cỡ mẫu của cuộc điều tra lần này tăng lên so với cuộc điều tra trước;

- Nếu tỷ lệ SSCM đó nhỏ hơn mức độ cho phép thì có thể điều chỉnh cỡ mẫu giảm đi.

Chú ý:

+ So sánh tỷ lệ SSCM là căn cứ quan trọng để điều chỉnh cỡ mẫu. Song đó không phải là căn cứ duy nhất, mà thực tế còn phải dựa vào một số yếu tố khác như sự thay đổi về quy mô tổng thể chung, thay đổi về số lượng chỉ tiêu điều tra,...

+ Điều kiện để áp dụng cách điều chỉnh cỡ mẫu trên đây là trong cuộc điều tra kỳ trước phải tính được tỷ lệ SSCM cho các chỉ tiêu chủ yếu.

Cách ước lượng này đơn giản và thuận tiện hơn nhiều so với cách tính cỡ mẫu theo lý thuyết, nhưng lại có cơ sở chắc chắn hơn so với cách xác định cỡ mẫu có tính chất ước đoán thuần túy theo kinh nghiệm.

d. Cách xác định cỡ mẫu chủ yếu dựa vào khả năng về kinh phí. Công thức xác định cỡ mẫu (n) trong trường hợp này như sau:

$$n = \frac{C - C_0}{Z} \quad ; \quad (1.1.12)$$

Trong đó:

C - Tổng kinh phí được cấp;

C_0 - Kinh phí chi cho các khâu chuẩn bị, tập huấn nghiệp vụ thu thập, xử lý và các chi phí chung khác;

Z - Chi phí cần thiết cho tất cả các khâu điều tra tính cho một đơn vị điều tra.

1.1.3.2. Phân bổ mẫu

Nếu địa bàn điều tra được chia thành các khu vực hoặc các tổ khác nhau và tiến hành điều tra trên tất cả các khu vực hoặc các tổ thì phải thực hiện phân bổ mẫu cho từng khu vực hoặc từng tổ đó.

Có nhiều cách phân bổ mẫu khác nhau, dưới đây chỉ giới thiệu một số cách phân bổ chủ yếu.

a. Phân bổ mẫu tỷ lệ thuận với quy mô tổng thể

Công thức xác định cỡ mẫu của từng tổ t (n_t) như sau:

$$n_t = \frac{N_t}{N} n = N_t f \quad ; \quad (1.1.13)$$

Trong đó:

t - Chỉ số thứ tự tổ ($t = 1, 2 \dots K$)

n - Số đơn vị mẫu chung;

n_t - Số đơn vị mẫu của tổ t;

N - Số đơn vị của tổng thể;

N_t - Số đơn vị của tổ t;

f - Tỷ lệ mẫu ($f = \frac{n}{N}$)

Các phân bổ mẫu tỷ lệ thuận với quy mô thường được áp dụng khi quy mô của các tổ tương đối đồng đều, phương sai và chi phí cho các tổ không khác nhau nhiều. Cách phân bổ này có ưu điểm: Dễ làm, không phải tính lại theo quyền số thực tế khi suy rộng kết quả là chỉ tiêu bình quân hoặc tỷ lệ cho tổng thể. Tuy nhiên, khi quy mô của các tổ khác nhau nhiều thì phân bổ tỷ lệ thuận với quy mô dễ làm cho các

tổ có quy mô nhỏ thường không đủ số lượng mẫu để đại diện cho tổ đó, ngược lại các tổ có quy mô lớn lại "thừa" cỡ mẫu. Mặt khác, việc tổ chức điều tra cũng như kinh phí cần thiết cho điều tra ở các tổ có quy mô lớn sẽ rất nặng nề, còn việc tổ chức điều tra cũng như kinh phí cần thiết cho điều tra ở các tổ có quy mô nhỏ lại quá nhẹ nhàng.

b. Phân bổ mẫu tỷ lệ với căn bậc hai của quy mô tổng thể

Công thức tính số đơn vị mẫu (n_t) của tổ t như sau:

$$n_t = n \cdot w_t \quad ; \quad (1.1.14a)$$

Trong đó:

n - Số đơn vị của tổng thể

w_t - Tỷ lệ giữa căn bậc hai số đơn vị của tổ t ($\sqrt{N_t}$) và tổng căn bậc hai số đơn vị của tất cả các tổ ($\sum_{t=1}^K \sqrt{N_t}$).

Như vậy công thức (1.1.14a) sẽ biến đổi như sau:

$$n_t = n \cdot w_t = n \left(\frac{\sqrt{N_t}}{\sum_{t=1}^K \sqrt{N_t}} \right) \quad ; \quad (1.1.14b)$$

Cách phân bổ này sẽ khắc phục nhược điểm của phân bổ tỷ lệ với quy mô tổng thể nhưng khi suy rộng phải tính lại theo quyền số thực tế.

c. Phân bổ Neyman

Phân bổ Neyman được coi là phân bổ tối ưu theo nghĩa thống kê thuần túy. Cỡ mẫu vừa tính theo tỷ lệ của quy mô, vừa tính đến sự khác nhau về độ biến động của chỉ tiêu nghiên cứu các tổ.

Công thức xác định cỡ mẫu (n_t) cho tổ t như sau:

$$n_t = n \cdot \frac{N_t S_t}{\sum_{t=1}^K N_t S_t} \quad \text{với } (t = 1, 2, \dots, K) \quad ; \quad (1.1.15)$$

Trong đó:

N_t - Tổng số đơn vị của tổ t ;

S_t - Độ lệch chuẩn của tổ thứ t .

Công thức trên cho thấy quy mô mẫu của các tổ tỷ lệ thuận với quy mô và phương sai của chúng. Tổ có phương sai lớn sẽ được phân nhiều đơn vị mẫu hơn tổ có phương sai nhỏ, tổ có quy mô lớn sẽ được phân nhiều đơn vị hơn các tổ có quy mô nhỏ.

d. Phân bổ mẫu tối ưu

Đây là cách phân bổ mẫu tối ưu đầy đủ hơn vì nó không những đề cập tới sự khác biệt về quy mô, sự biến động của chỉ tiêu được nghiên cứu giữa các tổ mà còn đề cập tới khả năng kinh phí của từng tổ. Công thức phân bổ mẫu tối ưu có dạng:

$$n_t = n \cdot \left(\frac{N_t S_t / \sqrt{c_t}}{\sum_{t=1}^K N_t S_t / \sqrt{c_t}} \right) \quad \text{với } t = 1, 2, \dots, K \quad ; \quad (1.1.16)$$

Trong đó: c_t - Chi phí điều tra cho tổ t .

Công thức trên cho thấy quy mô mẫu của các tổ tỷ lệ thuận với quy mô và phương sai của chúng. Mặt khác tỷ lệ nghịch với căn bậc hai của chi phí có thể có để thực hiện điều tra trên phạm vi của tổ. Vì vậy, phương pháp phân bổ mẫu này thường được áp dụng khi quy mô, phương sai và khả năng kinh phí của các tổ tương đối khác nhau.

e. Phân bổ mẫu có ưu tiên cho các tổ được đánh giá là quan trọng

Cách phân bổ mẫu này thường được áp dụng khi có sự khác nhau đáng kể giữa các tổ về hàm lượng thông tin cần thiết. Theo nguyên tắc này, các tổ có hàm lượng thông tin thấp được phân bổ cỡ mẫu nhỏ. Tư tưởng này thường ứng dụng trong điều tra các doanh nghiệp. Các doanh nghiệp thuộc tổ có quy mô lớn (có sản lượng hoặc số lượng

công nhân chiếm tỷ trọng lớn trong tổng sản lượng hoặc tổng số công nhân của các doanh nghiệp) thì phân bổ theo tỷ lệ mẫu lớn hơn. Ngược lại các doanh nghiệp có quy mô nhỏ hơn thì phân bổ tỷ lệ mẫu nhỏ hơn.

Tóm lại, phân bổ mẫu trong thực tế cần dựa vào việc phân tích đặc điểm cụ thể của các chỉ tiêu thống kê cần thu thập ở từng tổ. Mặc khác, cũng cần xét tới điều kiện thực tế diễn ra ở từng tổ. Điều này đặc biệt cần lưu ý trong khi phân bổ cỡ mẫu cho điều tra nhiều cấp.

1.1.3.3. Cách tính sai số chọn mẫu

Dưới đây sẽ trình bày công thức tính SSCM tương ứng với các phương pháp tổ chức chọn mẫu ngẫu nhiên đơn giản, mẫu phân tổ, mẫu 2 cấp và mẫu chùm

Cách trình bày công thức tính SSCM được bắt đầu từ một ví dụ giả định về danh sách các làng, bản với số hộ gia đình có vốn đầu tư cho sản xuất, kinh doanh (viết tắt là VĐT) của một địa bàn "Y" thuộc tỉnh miền núi (xem số liệu bảng 1.1).

Bảng 1.1. Danh sách những bản, làng với số hộ có đầu tư sản xuất, kinh doanh

TT bản	Tên bản	Số hộ	Vùng (*)	TT bản	Tên bản	Số hộ	Vùng (*)
1	A	9	1	11	N	10	2
2	I	10	2	12	E	13	1
3	D	11	3	13	P	11	3
4	B	11	1	14	F	11	2
5	K	12	1	15	G	12	1
6	Y	12	2	16	Q	9	3
7	C	9	3	17	Z	10	2

(*)Ghi chú: 1: Vùng cánh đồng; 2: Vùng khe dọc; 3: Vùng cao.

8	L	10	2	18	J	8	1
9	V	11	1	19	H	13	1
10	M	10	1	20	S	14	2
				Tổng số		216	

a. Phương pháp tổ chức chọn mẫu ngẫu nhiên đơn giản

* Tổ chức chọn mẫu

Khi tiến hành chọn mẫu ngẫu nhiên đơn giản chỉ việc lập danh sách các hộ gia đình có tên chủ hộ, địa chỉ và kèm theo số thứ tự từ 1 đến 216 của chung 20 làng, bản kể trên. Sau đó dùng bảng số ngẫu nhiên hoặc rút thăm chọn ngẫu nhiên không lặp lại từ danh sách được lập trong bảng để được số hộ cần điều tra (ở đây là chọn 20 hộ).

* Cách tính sai số chọn mẫu

Gọi i là số thứ tự của hộ gia đình trên địa bàn điều tra.

$i = 1, 2, \dots, N$ ($N = 216$ - Tổng số hộ của địa bàn điều tra)

$i = 1, 2, \dots, n$ ($n = 20$ - Số hộ chọn mẫu trên địa bàn)

x_i : Vốn đầu tư sản xuất, kinh doanh của hộ thứ i

Từ đó có công thức:

+ VĐT bình quân một hộ:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad ; \quad (1.1.17)$$

+ Phương sai mẫu:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad ; \quad (1.1.18)$$

+ Sai số chọn mẫu:

$$\mu = \sqrt{\frac{s^2}{n} \left(1 - \frac{n}{N}\right)} \quad ; \quad (1.1.19)$$

b. Phương pháp tổ chức chọn mẫu phân tổ

** Tổ chức chọn mẫu*

Trở lại ví dụ bảng 1.1 phân các bản thành 3 vùng địa hình, tức là 3 tổ (1: cánh đồng; 2: khe dọc; 3: vùng cao). Các vùng này có điều kiện kinh tế khác nhau và do đó có mức độ đầu tư cho sản xuất, kinh doanh của dân cư cũng khác nhau. Như vậy, việc phân chia các bản theo vùng địa hình sẽ liên quan nhiều đến VĐT cho SXKD của dân cư.

Gọi t là số thứ tự của các tổ ($t = 1, 2, \dots, K = 3$ - Số tổ của địa bàn điều tra);

Tổ 1: $t = 1$ (Vùng cánh đồng); Tổ 2: $t = 2$ (Vùng khe dọc);
Tổ 3: $t = 3$ (Vùng núi cao)

N_t - Số HGĐ của tổ (vùng) t

N - Tổng số hộ gia đình của địa bàn điều tra ($N = \sum_{t=1}^K N_t$)

n_t - Số hộ chọn mẫu của tổ (vùng) t

n - Tổng số hộ chọn mẫu của địa bàn ($n = \sum_{t=1}^K n_t$)

Cỡ mẫu mỗi tổ (n_t) có thể được chọn theo tỷ lệ đều nhau hoặc chọn không theo tỷ lệ đều nhau. Nếu chọn theo tỷ lệ đều nhau thì tỷ lệ chọn mẫu ở các tổ đều bằng f ($f = \frac{n}{N}$).

** Cách tính sai số chọn mẫu*

Gọi i là số thứ tự của HGĐ trong mỗi tổ

$i = 1, 2, \dots, N_t$ đối với tổng thể chung

$i = 1, 2, \dots, n_t$ đối với tổng thể mẫu

x_{it} - VĐT của hộ thứ i thuộc tổ t

Từ đó ta có công thức tính:

+ VĐT bình quân của các đơn vị thuộc tổ t:

$$\bar{x}_t = \frac{1}{n_t} \sum_{i=1}^{n_t} x_{it} \quad ; \quad (1.1.20)$$

+ VĐT bình quân của tất cả các đơn vị điều tra:

- Chọn theo tỷ lệ:

$$\bar{x} = \frac{1}{n} \sum_{t=1}^K \bar{x}_t n_t \quad ; \quad (1.1.21.a)$$

- Chọn không theo tỷ lệ:

$$\bar{x} = \frac{1}{N} \sum_{t=1}^K \bar{x}_t N_t \quad ; \quad (1.1.21.b)$$

+ Phương sai mẫu của các đơn vị trong tổ t:

$$s_t^2 = \frac{1}{n_t - 1} \sum_{i=1}^{n_t} (x_{it} - \bar{x}_t)^2 \quad ; \quad (1.1.22)$$

+ Sai số chọn mẫu:

- Chọn theo tỷ lệ:

$$\mu = \sqrt{\frac{\bar{s}_t^2}{n} \left(1 - \frac{n}{N}\right)} \quad ; \quad (1.1.23a)$$

$$\text{Trong đó: } \bar{s}_t^2 = \frac{\sum_{t=1}^K s_t^2 n_t}{\sum_{t=1}^K n_t}$$

- Chọn không theo tỷ lệ:

$$\mu = \frac{1}{N} \sqrt{\sum_{t=1}^K \frac{s_t^2}{n_t} \left(1 - \frac{n_t}{N_t}\right) N_t^2} \quad ; \quad (1.1.23b)$$

c. Phương pháp tổ chức chọn mẫu 2 cấp

* Tổ chức chọn mẫu

Cũng số liệu đã cho ở bảng 1.1 tiến hành chọn mẫu 2 cấp như sau: từ danh sách 20 làng bản chọn ngẫu nhiên không lặp lấy 4, tức là 20% số làng bản (chẳng hạn chọn được các bản số 1, 5, 12 và 19). Các bản được chọn là mẫu cấp I. Tiếp theo lập danh sách các HGĐ của 4 bản này, rồi từ các danh sách đó chọn ngẫu nhiên không lặp ra số hộ đều nhau cho mỗi bản (5 hộ) để tiến hành điều tra. Như vậy tổng số hộ được chọn là 20 (hộ là mẫu cấp II).

* Cách tính sai số chọn mẫu

Gọi j là số thứ tự của đơn vị mẫu cấp I (bản)

j = 1, 2, 3, ..., M (M = 20 - Tổng số bản của địa bàn điều tra)

j = 1, 2, 3, ..., m (m = 4 - Số bản được chọn vào mẫu cấp I)

i - Số thứ tự của đơn vị cấp II (HGĐ)

n - Tổng số đơn vị mẫu cấp II (HGĐ)

n* - Số đơn vị mẫu cấp II trong mỗi đơn vị mẫu cấp I (các đơn vị mẫu cấp I có số đơn vị mẫu cấp II bằng nhau: n* = n : m)

x_{ij} - Vốn đầu tư của HGĐ (đơn vị mẫu cấp II) thứ i thuộc bản (đơn vị mẫu cấp I) thứ j.

Ta có công thức tính:

+ VĐT bình quân của các đơn vị mẫu cấp II thuộc mẫu cấp I thứ j:

$$\bar{x}_j = \frac{1}{n^*} \sum_{i=1}^{n^*} x_{ij} \quad ; \quad (1.1.24)$$

+ VĐT bình quân của tất cả các đơn vị điều tra:

$$\bar{x} = \frac{1}{m} \sum_{j=1}^m \bar{x}_j = \frac{1}{n} \sum_{j=1}^m \sum_{i=1}^{n^*} x_{ij} \quad ; \quad (1.1.25)$$

+ Phương sai mẫu cấp II (hộ) thuộc từng đơn vị mẫu cấp I (bản) thứ j:

$$s_j^2 = \frac{1}{(n^* - 1)} \sum_{i=1}^{n^*} (x_{ij} - \bar{x}_j)^2 \quad ; \quad (1.1.26)$$

+ Bình quân các phương sai mẫu cấp II:

$$\bar{s}_j^2 = \frac{1}{m} \sum_{j=1}^m s_j^2 \quad ; \quad (1.1.27)$$

+ Phương sai mẫu cấp I:

$$s_b^2 = \frac{1}{m - 1} \sum_{j=1}^m (\bar{x}_j - \bar{x})^2 \quad ; \quad (1.1.28)$$

+ Sai số chọn mẫu:

$$\mu = \sqrt{\frac{s_b^2}{m} \left(1 - \frac{m}{M}\right) + \frac{\bar{s}_j^2}{m \cdot n^*} \left(1 - \frac{n^*}{N^*}\right)} \quad ; \quad (1.1.29)$$

Trong đó: Số đơn vị cấp II thực tế có bình quân trong mỗi đơn vị cấp I (N) : N* = N : M.

d. Phương pháp tổ chức chọn mẫu chùm

Trong mẫu chùm có hai loại: Mẫu chùm có kích thước bằng nhau và mẫu chùm có kích thước khác nhau. Sự khác nhau về kích thước của mẫu chùm liên quan đến sự khác nhau về cách tổ chức chọn mẫu và công thức tính các tham số chọn mẫu.

** Tổ chức chọn mẫu*

Tiếp tục nghiên cứu ví dụ 1.1. Nếu xác định chùm là một bản và cũng tiến hành điều tra cỡ mẫu $n = 20$ hộ gia đình thì cách tiến hành như sau:

+ Với cỡ mẫu có kích thước các chùm bằng nhau (do người tổ chức điều tra ấn định) thì số chùm (m) cần chọn được xác định bằng cách chia tổng số mẫu cần điều tra (n) cho số mẫu qui định trong một chùm (n^*), tức là $n : n^* = m$. Cũng với ví dụ trên, cần điều tra 20 hộ ($n = 20$) và giả sử qui định mỗi chùm chọn 10 hộ ($n^* = 10$) thì số chùm (bản) phải điều tra: $m = 20 : 10 = 2$ chùm.

Sau khi xác định được số chùm cần chọn, ta lập danh sách tất cả các chùm rồi chọn ngẫu nhiên không lặp lại từ danh sách đã cho 2 chùm (bản) để tiến hành điều tra thực tế các đơn vị thuộc các chùm đó.

+ Với cỡ mẫu có kích thước các chùm khác nhau thì quá trình chọn mẫu được tiến hành qua các bước sau đây:

- Chia tổng số HGĐ của địa bàn điều tra cho số bản để xác định số hộ bình quân có trong một chùm:

$$N^* = 216 : 20 \approx 11$$

- Chia số mẫu (HGĐ) cần chọn cho số hộ có trong một chùm để xác định số chùm cần điều tra (m):

$$m = 20 : 11 \approx 2 \text{ chùm}$$

Trên cơ sở danh sách các bản ở bảng 1.1, tiến hành chọn 2 chùm, rồi tổ chức điều tra thực tế toàn bộ số HGĐ của 2 chùm đó.

Khi chọn mẫu chùm có kích thước khác nhau để điều tra sẽ có những trường hợp sau đây:

- Nếu ở 2 chùm có vừa đủ 20 HGĐ thì điều tra hết 20 hộ.
- Nếu ở 2 chùm có số HGĐ lớn hơn ($>$)20 thì điều tra hết 20 hộ2,

số dư ra bỏ lại không điều tra tiếp.

- Nếu ở 2 chùm có số HGĐ nhỏ hơn ($<$)20 thì điều tra hết số 2GD của 2 bản đã chọn. Sau chọn thêm một bản thứ ba trong số 18 bản còn lại và điều tra thêm số hộ cho đủ 20.

** Cách tính sai số chọn mẫu*

Gọi j là thứ tự các chùm (bản), ở đây: $j = 1, 2, 3, \dots, M$ ($M = 20$ - toàn bộ số bản có trong địa bàn điều tra) và $j = 1, 2, 3, \dots, m$ ($m = 2$ - số chùm chọn mẫu).

Gọi i là số thứ tự của HGĐ, ở đây $i = 1, 2, 3, \dots, n_j$ (n_j là số hộ có của một chùm - bản).

$$\text{Trong đó: } \sum_{j=1}^m n_j = n \text{ (n là số mẫu điều tra)}$$

Nếu chọn mẫu chùm có kích thước bằng nhau thì các n_j bằng nhau và bằng n^* (n^* là số đơn vị trong một chùm)

Gọi x_{ij} : VĐT của hộ thứ i thuộc chùm j

Ta có công thức tính cho hai trường hợp:

+ Chùm có kích thước bằng nhau:

- VĐT bình quân của các đơn vị trong mỗi chùm thứ j

$$\bar{x}_j = \frac{1}{n^*} \sum_{i=1}^{n^*} x_{ij} \quad ; \quad (1.1.30)$$

- VĐT bình quân của tất cả các đơn vị điều tra

$$\bar{x} = \frac{1}{m} \sum_{j=1}^m \bar{x}_j \quad ; \quad (1.1.31)$$

- Phương sai giữa các chùm

$$s_b^2 = \frac{1}{m-1} \sum_{j=1}^m (\bar{x}_j - \bar{x})^2 \quad ; \quad (1.1.32)$$

- Sai số chọn mẫu

$$\mu = \sqrt{\frac{s_b^2}{m} \left(1 - \frac{m}{M}\right)} \quad ; \quad (1.1.33)$$

+ Chùm có kích thước khác nhau:

- VĐT bình quân của các đơn vị trong mỗi chùm thứ j

$$\bar{x}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} x_{ij} \quad ; \quad (1.1.34)$$

- VĐT bình quân của tất cả các đơn vị điều tra

$$\bar{x} = \frac{\sum_{j=1}^m \bar{x}_j n_j}{\sum_{j=1}^m n_j} = \frac{1}{n} \sum_{j=1}^m \sum_{i=1}^{n_j} x_{ij} \quad ; \quad (1.1.35)$$

- Phương sai giữa các chùm:

$$s_b^2 = \frac{1}{\left(n - \frac{n}{m}\right)} \sum_{j=1}^m (\bar{x}_j - \bar{x})^2 n_j \quad ; \quad (1.1.36)$$

- Sai số chọn mẫu: Như công thức 1.1.33.

1.2. SAI SỐ TRONG ĐIỀU TRA THỐNG KÊ

Trong điều tra thống kê có hai loại sai số: Sai số chọn mẫu (sai số do tính đại diện của số liệu vì chỉ chọn một bộ phận các đơn vị để điều tra) và sai số phi chọn mẫu (sai số thuộc về lỗi của các quy định, hướng dẫn, giải thích tài liệu điều tra, do sai sót của việc cân đong, đo

đếm, cung cấp thông tin, ghi chép, đánh mã, nhập tin,...) từ đây gọi là "sai số điều tra".

Sai số chọn mẫu (SSCM) chỉ phát sinh trong điều tra chọn mẫu khi tiến hành thu thập ở một bộ phận các đơn vị tổng thể (gọi là mẫu) rồi dùng kết quả suy rộng cho toàn bộ tổng thể chung. SSCM phụ thuộc vào cỡ mẫu (mẫu càng lớn thì sai số càng nhỏ), vào độ đồng đều của chỉ tiêu nghiên cứu (độ đồng đều cao thì sai số chọn mẫu càng nhỏ) và phương pháp tổ chức điều tra chọn mẫu. Còn sai số điều tra xảy ra cả trong điều tra chọn mẫu và điều tra toàn bộ.

Trong thực tế công tác điều tra thống kê hiện nay, phương pháp chọn mẫu được áp dụng ngày càng nhiều và có hiệu quả. Số liệu thu được từ điều tra chọn mẫu ngày càng phong phú, đa dạng và phục vụ kịp thời các yêu cầu sử dụng. Bên cạnh đó chất lượng số liệu của điều tra chọn mẫu cũng còn những hạn chế nhất định. Có một số ý kiến hiện nay đánh giá không công bằng và thiếu khách quan về kết quả điều tra chọn mẫu, cho rằng số liệu chưa sát với thực tế vì chỉ điều tra một bộ phận rồi suy rộng cho tổng thể.

Tất nhiên cũng phải thấy rằng đã là điều tra chọn mẫu thì không thể tránh khỏi sai số chọn mẫu nhưng mức độ sai số chọn mẫu của phần lớn những chỉ tiêu trong các cuộc điều tra thống kê hiện nay thường là ở phạm vi cho phép nên chấp nhận được. Hơn nữa khi cần thiết ta có thể chủ động giảm được sai số chọn mẫu bằng cách điều chỉnh cỡ mẫu và tổ chức chọn mẫu một cách khoa học, tuân thủ đúng nguyên tắc chọn mẫu.

Điều đáng nói và cần quan tâm hơn trong điều tra thống kê chính là sai số phi chọn mẫu. Loại sai số này xảy ra ở cả ba giai đoạn điều tra, liên quan đến tất cả các đối tượng tham gia điều tra thống kê và ảnh hưởng đáng kể đến chất lượng số liệu thống kê.

Dưới đây sẽ đi sâu nghiên cứu về sai số phi chọn mẫu - sai số điều tra, xảy ra trong cả ba giai đoạn nhưng chỉ đề cập đến sai số liên quan tới những công việc, những đối tượng thường gặp nhiều hơn.

1.2.1. Sai số trong quá trình chuẩn bị điều tra thống kê

Trong công tác điều tra thống kê, chuẩn bị điều tra giữ một vai trò cực kỳ quan trọng. Chất lượng của khâu chuẩn bị điều tra sẽ ảnh hưởng cả đến quá trình thu thập số liệu và cuối cùng là đến chất lượng của số liệu điều tra. Một cuộc điều tra được chuẩn bị kỹ lưỡng, chu đáo và đầy đủ sẽ là cơ sở đầu tiên để giảm sai số điều tra nhằm nâng cao chất lượng của số liệu thống kê.

a. Sai số điều tra liên quan tới việc xác định mục đích, nội dung và đối tượng điều tra

Xác định mục đích điều tra là làm rõ yêu cầu của cuộc điều tra phải trả lời những câu hỏi gì, đạt được những mục tiêu nào của công tác quản lý. Yêu cầu của mục đích điều tra phải rõ ràng, dứt khoát và đó chính là căn cứ để xác định nội dung cũng như đối tượng điều tra một cách đúng đắn, đầy đủ, phù hợp, không bị chệch hướng.

Cùng một đơn vị điều tra, nếu có mục đích điều tra khác nhau với cách tiếp cận thu thập thông tin khác nhau thì sẽ có nội dung cũng như đối tượng điều tra khác nhau.

Xác định đúng nội dung và đối tượng điều tra, một mặt làm cho số liệu thu thập được sẽ đáp ứng những yêu cầu sử dụng, số liệu đảm bảo "vừa đủ". Mặt khác, xác định đúng nội dung và đối tượng điều tra là cơ sở để thiết kế bảng hỏi một cách khoa học và có điều kiện thuận lợi để tiếp cận với đối tượng cung cấp thông tin, đảm bảo thông tin thu được phù hợp và phản ánh đúng thực tế khách quan.

Tóm lại việc xác định đúng mục đích, nội dung và đối tượng điều tra làm cho cuộc điều tra thực hiện đúng hướng, đúng yêu cầu là một trong những điều kiện tiên quyết để đảm bảo chất lượng số liệu, giảm sai số trong điều tra thống kê.

b. Sai số liên quan tới việc xây dựng các khái niệm, định nghĩa dùng trong điều tra

Khái niệm, định nghĩa dùng trong điều tra giúp cho hiểu rõ nội dung, bản chất cũng như phạm vi xác định thông tin của số liệu thống kê cần thu thập.

Như ta đã biết thống kê nghiên cứu mặt lượng trong quan hệ mật thiết với mặt chất của hiện tượng kinh tế - xã hội số lớn. Chính các khái niệm, định nghĩa là phản ánh về mặt chất của hiện tượng, là cơ sở để nhận biết, phân biệt hiện tượng này với hiện tượng khác cũng như xác định phạm vi của hiện tượng nghiên cứu. Nếu khái niệm, định nghĩa chuẩn xác, rõ ràng, được giải thích đầy đủ, cặn kẽ là cơ sở để xác định và thu thập số liệu thống kê phản ánh đúng thực tế khách quan. Ngược lại nếu khái niệm, định nghĩa không đúng, mập mờ, thiếu rõ ràng thì việc xác định, đo tính (lượng hoá) hiện tượng sẽ bị sai lệch.

Ví dụ: Khi điều tra cán bộ khoa học công nghệ có trình độ "trên đại học", xét về chất, trên đại học phải là những người đã tốt nghiệp và có bằng thạc sĩ, tiến sĩ và tiến sĩ khoa học. Trong thực tế có cuộc điều tra thống kê ở nước ta chỉ đưa ra khái niệm "trên đại học" chung chung, thiếu cụ thể. Điều này làm cho những người tham gia điều tra (kể cả điều tra viên lẫn đối tượng trả lời) hiểu khái niệm cán bộ khoa học công nghệ có trình độ trên đại học rất khác nhau. Một số ít người đã hiểu đúng với nghĩa trình độ trên đại học phải gồm những người có bằng thạc sĩ, tiến sĩ và tiến sĩ khoa học; phần đông còn lại đã hiểu không đúng và cho là trên đại học gồm những người đã tốt nghiệp đại học sau đó được đi thực tập sinh sau đại học và thậm chí còn cả những người đã tốt nghiệp đại học nhưng chỉ được đi tập trung để đào tạo bồi dưỡng thêm về nghiệp vụ một vài tháng.

Thực tế này đã làm cho số liệu điều tra được về cán bộ khoa học công nghệ có trình độ "trên đại học" tăng lên hơn hai lần so với số thực tế có tại thời điểm điều tra.

Như vậy, những lỗi trong việc xây dựng các khái niệm, định nghĩa và nội dung thông tin về tiêu thức, chỉ tiêu thống kê sẽ ảnh hưởng trực tiếp đến chất lượng số liệu thống kê. Đây là hiện tượng

khá phổ biến trong điều tra thống kê ở nước ta hiện nay.

Để có số liệu tốt, giảm bớt sai số điều tra, một vấn đề có tính chất nguyên tắc đó là phải chuẩn hoá các khái niệm, định nghĩa về các tiêu thức, chỉ tiêu của điều tra thống kê. Đồng thời phải giải thích rõ ràng, đầy đủ và cụ thể hoá các khái niệm, định nghĩa cho phù hợp với từng cuộc điều tra riêng biệt.

c. Sai số điều tra liên quan tới thiết kế bảng hỏi, xây dựng các bảng danh mục và mã số dùng trong điều tra

Trong điều tra thống kê, bảng hỏi là vật mang tin, là công cụ giúp điều tra viên điền thông tin hoặc đánh dấu, đánh mã vào các ô, dòng, cột phù hợp theo nội dung trả lời của các câu hỏi tương ứng với các tiêu thức ghi ở bảng hỏi dùng trong điều tra.

Nếu các câu hỏi phức tạp, khó hiểu, khó trả lời, khó xác định hoặc khó điền thông tin thì khi đó thông tin thu được sẽ kém chính xác, không đáp ứng yêu cầu của số liệu điều tra.

Cùng với bảng hỏi, các bảng danh mục và các mã số có vai trò quan trọng trong quá trình tổng hợp số liệu thống kê. Thông tin thu được dù đảm bảo độ tin cậy cần thiết, nhưng nếu bảng danh mục dùng cho điều tra không chuẩn xác, các mã số không rõ ràng, khó áp dụng dẫn tới việc đánh sai, đánh nhầm và tất nhiên như vậy số liệu tổng hợp sẽ bị sai lệch.

Để giảm sai số điều tra, bảng hỏi phải được thiết kế một cách khoa học, đáp ứng đầy đủ nhu cầu thông tin theo nội dung điều tra đã được xác định, bảo đảm mối liên hệ logic và tính thống nhất giữa các câu hỏi. Mặt khác, các câu hỏi phải đơn giản, dễ hiểu, dễ trả lời, dễ ghi chép, phù hợp với trình độ của điều tra viên và đặc điểm về nguồn thông tin của từng loại câu hỏi. Thiết kế bảng hỏi còn phải đảm bảo thuận lợi cho việc áp dụng công nghệ thông tin. Các bảng danh mục phải có nội dung phù hợp với những thông tin cần thu thập và được mã hoá một cách khoa học theo yêu cầu tổng hợp của điều tra. Danh mục vừa phải phù hợp với yêu cầu của từng cuộc điều tra, vừa phải

đáp ứng và thống nhất với danh mục phục vụ cho tổng hợp chung của công tác thống kê. Nội dung bảng danh mục và cách mã hoá phải được giải thích đầy đủ và hướng dẫn cụ thể.

d. Sai số điều tra liên quan tới việc lựa chọn điều tra viên và hướng dẫn nghiệp vụ

Điều tra viên là người trực tiếp truyền đạt mục đích, nội dung, yêu cầu điều tra đến các đối tượng cung cấp thông tin, đồng thời trực tiếp phỏng vấn, lựa chọn thông tin để ghi vào bảng hỏi (nếu là điều tra trực tiếp). Vì vậy, điều tra viên có vai trò rất quan trọng trong việc đảm bảo chất lượng số liệu trong điều tra.

Nếu điều tra viên không nắm vững mục đích của cuộc điều tra, không hiểu hết nội dung thông tin cần thu thập thì sẽ truyền đạt không đúng các yêu cầu cần thiết cho đối tượng trả lời. Ngay cả khi điều tra viên nắm được nghiệp vụ, nhưng nếu thiếu ý thức trách nhiệm, chỉ phỏng vấn và ghi chép cho xong việc, hoặc cách tiếp cận với đối tượng điều tra không tốt thì cũng sẽ dẫn đến kết quả số liệu điều tra thu được không theo ý muốn.

Như vậy, việc lựa chọn điều tra viên không tốt cũng là nguyên nhân không kém phần quan trọng làm cho sai số điều tra tăng lên, ảnh hưởng đến chất lượng số liệu. Vì vậy, muốn giảm bớt loại sai số điều tra này, cần tuyển chọn điều tra viên có trình độ nhất định, nắm được nghiệp vụ, có kinh nghiệm thực tế về điều tra thống kê, đồng thời phải có ý thức và tinh thần trách nhiệm cao.

Sau khi lựa chọn được điều tra viên cần tổ chức tập huấn nghiệp vụ đầy đủ và thống nhất. Trong lớp tập huấn bên cạnh giải thích biểu mẫu điều tra cần cung cấp thêm những kiến thức về xã hội, phổ biến những kinh nghiệm thực tế và cách tiếp cận đối tượng điều tra, cách ứng xử trong thực tế. Đối với các cuộc điều tra thống kê có nội dung phức tạp và quy mô lớn, cần tiến hành điều tra thử để kịp thời rút kinh nghiệm, đảm bảo hướng dẫn nghiệp vụ gắn với điều tra thực địa.

Trong điều tra chọn mẫu, khi hướng dẫn nghiệp vụ cần chỉ rõ lộ

trình điều tra theo từng cấp chọn mẫu, xác định địa bàn điều tra, lập danh sách địa bàn và đối tượng điều tra chọn mẫu (có địa chỉ cụ thể), quy định rõ những trường hợp mất mẫu phải thay đổi như thế nào, thay đổi đến đâu để tránh tình trạng điều tra viên thay đổi mẫu tùy tiện theo ý chủ quan của họ, v.v...

1.2.2. Sai số trong quá trình tổ chức điều tra

a. Sai số điều tra liên quan đến quan hệ giữa yêu cầu về nội dung thông tin và quỹ thời gian, các điều kiện vật chất cần cho thu thập số liệu

Nếu trong các cuộc điều tra thống kê phải thu thập quá nhiều chỉ tiêu có nội dung thông tin phức tạp, tốn nhiều thời gian để giải thích, phỏng vấn và ghi chép; trong khi đó quỹ thời gian và kinh phí dành cho công việc này lại không tương xứng, làm cho điều tra viên không đủ điều kiện để tiếp cận tìm hiểu tình hình thực tế, giải thích một cách đầy đủ, cặn kẽ về mục đích, yêu cầu và nội dung điều tra... cho người cung cấp thông tin thì có thể họ sẽ không khai báo, hoặc khai báo qua loa, sai với thực tế. Đặc biệt có những loại thông tin phải hỏi tường thì càng không đủ thời gian để nhớ lại... Tất cả những điều đó làm cho số liệu thu thập được sai số nhiều, không phản ánh đúng thực tế khách quan.

Để nâng cao chất lượng số liệu thống kê, giảm sai số khi tổ chức điều tra, phải cân đối giữa nhu cầu thu thập thông tin với khả năng về điều kiện kinh phí và quỹ thời gian dành cho điều tra. Không nên tổ chức một cuộc điều tra đòi hỏi thu thập quá nhiều chỉ tiêu; đặc biệt phải giới hạn những chỉ tiêu thu thập quá khó và tính toán phức tạp. Hơn nữa tùy thuộc vào đặc điểm và nội dung thông tin của các chỉ tiêu khác nhau, thuộc các đối tượng khác nhau để có cách tiếp cận thu thập thông tin cho hợp lý. Có thể chỉ tiêu này cần thu thập từ những nội dung chi tiết rồi tổng hợp chung lại, nhưng chỉ tiêu kia chỉ cần lấy số liệu khái quát. Không nên cho rằng bất kỳ chỉ tiêu nào, nội dung thông tin nào cũng phải lấy từ số liệu chi tiết mới là chính xác.

b. Sai số điều tra liên quan đến điều tra viên

Như trên đã nói để nâng cao chất lượng số liệu, giảm sai số điều tra, một trong những yêu cầu là phải chọn những người điều tra đủ tiêu chuẩn về chuyên môn và tinh thần trách nhiệm.

Ngoài những yêu cầu trên, điều tra viên khi được phân công về địa bàn điều tra, còn đòi hỏi phải làm quen với địa bàn, tìm hiểu thực tế về phong tục, tập quán, về điều kiện đi lại, sinh hoạt của địa phương.

Khi điều tra, điều tra viên phải kết hợp được kiến thức chuyên môn về điều tra đã được hướng dẫn với tình hình thực tế ở địa bàn điều tra, vừa phải giữ đúng nguyên tắc quy định cho điều tra, vừa phải có được những xử lý linh hoạt và hài hoà. Phần lớn những thắc mắc của đối tượng điều tra, điều tra viên phải tự mình tìm ra hướng giải đáp. Chỉ những trường hợp cần thiết mới ghi lại để xin ý kiến về cách xử lý của cấp chỉ đạo cao hơn.

c. Sai số điều tra liên quan đến ý thức, tâm lý và khả năng hiểu biết của người trả lời

Ở đây việc trả lời câu hỏi có thể không tốt do ba nguyên nhân thuộc người cung cấp thông tin như sau:

- Về ý thức của người trả lời: Nếu họ không có tinh thần trách nhiệm cao, cho là cung cấp thông tin thế nào cũng được, nói cho xong việc thì có thể khi điều tra, người cung cấp thông tin sẽ lấy lý do này, lý do khác để không trả lời hoặc trả lời không hết, không đúng sự thật. Không ít trường hợp người trả lời còn cố tình khai không đúng vì lợi ích kinh tế và mục đích khác.

- Về tâm lý, nhiều người cung cấp thông tin không muốn trả lời những câu hỏi liên quan đến đời tư, đến mức sống, đến sự bí mật kín đáo của họ, của đơn vị họ. Ví dụ, khi điều tra thu thập thông tin mức thu nhập của hộ gia đình, phần lớn các chủ hộ nhất là những người có thu nhập cao thường không muốn nói thật, nói hết mức thu nhập của

mình. Một ví dụ khác một người phụ nữ đi nạo thai trong trường hợp giấu gia đình họ sẽ không muốn khai vì không muốn cho những người thân trong gia đình biết đến.

- Về nhận thức của người trả lời, nhiều người do nhận thức có hạn, không thấy rõ được mục đích, yêu cầu điều tra, không hiểu được nội dung câu trả lời... do vậy họ không thể trả lời hoặc trả lời không đúng với yêu cầu câu hỏi.

Qua đây cho thấy, để giảm bớt sai số điều tra, điều tra viên phải có cách tiếp cận hợp lý với từng loại đối tượng điều tra, ngoài kiến thức chuyên môn còn phải hiểu biết về xã hội, giải thích cho người được phỏng vấn về mục đích, ý nghĩa, về nguyên tắc cung cấp và bảo mật thông tin riêng, về trách nhiệm và quyền hạn của người cung cấp thông tin, giải thích cho họ hiểu nội dung câu hỏi một cách thuận tiện nhất, gợi ý cho họ những cách trả lời để đi đến có được số liệu thật.

d. Sai số điều tra liên quan đến các phương tiện cân, đo, đong, đo lường

Tất cả các khâu khác chuẩn bị tốt, nhưng nếu các loại phương tiện như cân, thước đo, dụng cụ đo huyết áp... dùng cho các chỉ tiêu phải thực hiện kiểm tra, đo, đếm trực tiếp mà không được chuẩn bị tốt thì cũng sẽ sai sót dẫn đến sai số trong điều tra. Ví dụ, điều tra để xác định mức độ suy dinh dưỡng của trẻ em. Nếu ta dùng loại cân không chuẩn thì sẽ cân không chính xác, dẫn đến số liệu tổng hợp về tỷ lệ trẻ em suy dinh dưỡng sẽ không đúng, hoặc là cao hơn, hoặc là thấp hơn thực tế.

Như vậy, việc chuẩn bị tốt các phương tiện đo lường khi điều tra cũng là biện pháp cần thiết để giảm sai số điều tra.

1.2.3. Sai số liên quan đến quá trình xử lý thông tin

Sai số điều tra còn có thể xảy ra vì sai sót trong khâu đánh mã, nhập tin trong quá trình tổng hợp, xử lý số liệu.

Số liệu thu về phải được kiểm tra sơ bộ trước khi đánh mã, nhập

tin. Việc kiểm tra này có thể phát hiện ra những trường hợp hiểu đúng nhưng ghi chép sai như nhầm đơn vị tính: Cái ghi sai thành 1000 cái, 1 đồng thành 1000 đồng; điền sai vị trí của thông tin.... Bảng kinh nghiệm nghề nghiệp cũng như quan hệ logic tính toán giữa các câu hỏi, người kiểm tra có thể phát hiện được những loại sai sót kiểu này. Kiểm tra sơ bộ còn có thể phát hiện những trường hợp có "số liệu lạ" (quá cao hoặc quá thấp so với mức bình quân chung). Những loại sai sót trên đây nhân viên kinh tế có thể tự sửa hoặc nếu trong những trường hợp cần thiết phải kiểm tra xác minh lại. Làm tốt khâu kiểm tra sơ bộ cũng là công việc góp phần quan trọng để giảm sai số điều tra.

Sau kiểm tra sơ bộ là công đoạn đánh mã và nhập tin. Số liệu ghi đúng, ghi đầy đủ được kiểm tra kỹ lưỡng, nhưng nếu đánh mã sai, hoặc nhập tin sai thì cũng dẫn đến kết quả tổng hợp sai.

Sai sót trong đánh mã có thể là lựa chọn mã không phù hợp với nội dung của thông tin (hoặc là do bảng mã không cụ thể, khó xác định, hoặc là khả năng liên hệ vận dụng mã của người đánh mã không tốt), đánh mã sai (mã này lẫn với mã kia) hoặc có mã đúng nhưng lộn số (ví dụ 51 thành 15), v.v...

Để khắc phục sai sót trong khâu đánh mã, trước hết phải có bảng mã tốt, cụ thể, phù hợp với nội dung thông tin cần thu thập. Bên cạnh những mã cụ thể cần có những mã chung để cho người đánh mã có cơ sở vận dụng cho những trường hợp thực tế xảy ra nhưng chưa có mã trong danh mục mã cụ thể (gọi là các trường hợp khác). Mặt khác, người đánh mã phải được hướng dẫn đầy đủ về yêu cầu, nguyên tắc và kỹ thuật đánh mã, khi thực hiện phải biết vận dụng và xử lý linh hoạt nhưng tuyệt đối không được tùy tiện, người đánh mã còn kết hợp chặt chẽ với các bộ phận khác trong cùng khâu tổng hợp, xử lý số liệu.

Sau đánh mã là khâu nhập tin và khâu này cũng thường xuyên xảy ra sai sót. Loại sai sót này thường xảy ra trong các trường hợp sau: Nhập tin đúng hoặc bỏ qua không nhập tin, nhập mã sai, lộn số, v.v...

Để khắc phục những sai sót khi nhập tin, trước hết phải lựa chọn những nhân viên nhập tin có khả năng nhập tốt, ít nhầm lẫn, có tinh thần trách nhiệm cao, tuân thủ nghiêm túc những quy trình và nguyên tắc nhập tin đã được hướng dẫn thống nhất.

Trên góc độ công nghệ thông tin, phải có chương trình nhập hợp lý, khoa học, có được những lệnh cho phép tự kiểm tra để phát hiện những lỗi nhập tin.

Trong nhiều trường hợp phải phân công chéo để nhập tin hai lần rồi so sánh đối chiếu số liệu nhập để tìm ra những trường hợp không thống nhất thuộc về lỗi nhập tin.

Đối với các cuộc điều tra thống kê thực tế hiện nay, những lỗi nhập tin ảnh hưởng đến sai số điều tra không phải là nhỏ. Tuy nhiên, sai số do lỗi nhập tin hoàn toàn có điều kiện để khắc phục tốt.

PHẦN HAI

BIỂU HIỆN CÁC MỨC ĐỘ CỦA HIỆN TƯỢNG KINH TẾ - XÃ HỘI

Nghiên cứu các mức độ của hiện tượng kinh tế - xã hội là yêu cầu quan trọng của việc tổng hợp, tính toán và phân tích thống kê nhằm biểu hiện mặt lượng trong quan hệ mật thiết với mặt chất của hiện tượng nghiên cứu trong điều kiện thời gian và không gian cụ thể nhờ vào sự trợ giúp của các phương pháp thống kê.

Để biểu hiện các mức độ của hiện tượng trong thống kê dùng các số tuyệt đối (phản ánh quy mô), các số tương đối (phản ánh tốc độ, quan hệ tỷ lệ, cơ cấu, trình độ phổ biến), các số bình quân (phản ánh mức độ điển hình); toàn cục, phương sai, độ lệch chuẩn, hệ số biến thiên (phản ánh độ biến động của tiêu thức); đường cong Lorenz, hệ số GINI (phản ánh mức độ tập trung hay phân tán của phân phối),...

Dưới đây là nội dung, phương pháp tính và điều kiện vận dụng của các đại lượng đó.

2.1. SỐ TUYỆT ĐỐI (TRONG THỐNG KÊ)

Số tuyệt đối là chỉ tiêu biểu hiện quy mô, khối lượng của hiện tượng hoặc quá trình kinh tế - xã hội trong điều kiện thời gian và không gian cụ thể.

Số tuyệt đối trong thống kê bao gồm các con số phản ánh quy mô của tổng thể hay của từng bộ phận trong tổng thể (số doanh nghiệp, số nhân khẩu, số học sinh đi học, số lượng cán bộ khoa học,...) hoặc tổng các trị số theo một tiêu thức nào đó (tiền lương của công nhân, giá trị sản xuất công nghiệp, tổng sản phẩm trong nước (GDP), v.v...).

Số tuyệt đối trong thống kê bao giờ cũng có đơn vị tính cụ thể. Đơn vị tính số tuyệt đối có thể là đơn vị hiện vật tự nhiên (cái, con,

chiếc, kg, mét, v.v...), đơn vị hiện vật quy ước tức là đơn vị quy đổi theo một tiêu chuẩn nào đó (nước mắm quy theo độ đậm; than quy theo hàm lượng calo; xà phòng quy theo tỷ lệ chất béo; vải quy theo mét độ dài tiêu chuẩn,...), đơn vị tiền tệ (đồng, nhân dân tệ, đô la v.v...),... đơn vị thời gian (giờ, ngày, tháng, năm) và đơn vị kép (tấn-km, ngày -người,...).

Số tuyệt đối dùng để đánh giá và phân tích thống kê, là căn cứ không thể thiếu được trong việc xây dựng chiến lược phát triển kinh tế, tính toán các mặt cân đối, nghiên cứu các mối quan hệ kinh tế - xã hội, là cơ sở để tính toán các chỉ tiêu tương đối và bình quân.

Có hai loại số tuyệt đối: Số tuyệt đối thời kỳ và số tuyệt đối thời điểm.

Số tuyệt đối thời kỳ: Phản ánh quy mô, khối lượng của hiện tượng trong một thời kỳ nhất định. Ví dụ: Giá trị sản xuất công nghiệp trong 1 tháng, quý hoặc năm. Sản lượng lương thực năm 2002, năm 2003, năm 2004,...

Số tuyệt đối thời điểm: Phản ánh quy mô, khối lượng của hiện tượng ở một thời điểm nhất định như: dân số của một địa phương nào đó có đến 0 giờ ngày 01/04/1999; giá trị tài sản cố định có đến 31/12/2003; lao động làm việc của doanh nghiệp vào thời điểm 1/7/2004,...

2.2. SỐ TƯƠNG ĐỐI (TRONG THỐNG KÊ)

Số tương đối là chỉ tiêu biểu hiện quan hệ so sánh giữa hai chỉ tiêu thống kê cùng loại nhưng khác nhau về thời gian hoặc không gian, hoặc giữa hai chỉ tiêu khác loại nhưng có quan hệ với nhau. Trong hai chỉ tiêu để so sánh của số tương đối, sẽ có một số được chọn làm gốc (chuẩn) để so sánh.

Số tương đối có thể được biểu hiện bằng số lần, số phần trăm (%) hoặc phần nghìn h (‰), hay bằng các đơn vị kép h (người /km², người /1000 người; đồng /1000đồng,...).

Ví dụ: So với năm 2001, GDP năm 2002 của Việt Nam bằng 1, 07 lần hoặc 107,0%; tỷ lệ dân số thành thị của cả nước năm 2002 là 25,1%, mật độ dân số của Việt Nam năm 2002 là 239 người /km²,...

Trong công tác thống kê, số tương đối được sử dụng rộng rãi để phản ánh những đặc điểm về kết cấu, quan hệ tỷ lệ, trình độ phát triển, trình độ hoàn thành kế hoạch, trình độ phổ biến của hiện tượng kinh tế - xã hội được nghiên cứu trong điều kiện thời gian và không gian nhất định.

Số tương đối phải được vận dụng kết hợp với số tuyệt đối. Số tương đối thường là kết quả của việc so sánh giữa hai số tuyệt đối. Số tương đối tính ra có thể rất khác nhau, tùy thuộc vào việc lựa chọn gốc so sánh. Có khi số tương đối có giá trị rất lớn nhưng ý nghĩa của nó không đáng kể vì trị số tuyệt đối tương ứng của nó lại rất nhỏ. Ngược lại, có số tương đối tính ra khá nhỏ nhưng lại mang ý nghĩa quan trọng vì trị số tuyệt đối tương ứng của nó có quy mô đáng kể. Ví dụ: 1% dân số Việt Nam tăng lên trong những năm 1960 đồng nghĩa với dân số tăng thêm 300 nghìn người, nhưng 1% dân số tăng lên trong những năm 2000 lại đồng nghĩa với dân số tăng thêm 800 nghìn người.

Căn cứ vào nội dung mà số tương đối phản ánh, có thể phân biệt: số tương đối động thái, số tương đối kế hoạch, số tương đối kết cấu, số tương đối cường độ, và số tương đối không gian.

2.2.1. Số tương đối động thái

Số tương đối động thái là chỉ tiêu phản ánh biến động theo thời gian về mức độ của chỉ tiêu kinh tế - xã hội. Số tương đối này tính được bằng cách so sánh hai mức độ của chỉ tiêu được nghiên cứu ở hai thời gian khác nhau. Mức độ của thời kỳ được tiến hành nghiên cứu thường gọi là mức độ của kỳ báo cáo, còn mức độ của một thời kỳ nào đó được dùng làm cơ sở so sánh thường gọi là mức độ kỳ gốc. Ví dụ: So với năm 2001, GDP năm 2002 của Việt Nam bằng 1, 07 lần hoặc 107,0%.

2.2.2. Số tương đối so sánh

Số tương đối so sánh là chỉ tiêu phản ánh quan hệ so sánh giữa hai bộ phận trong một tổng thể, hoặc giữa hai hiện tượng cùng loại nhưng khác nhau về điều kiện không gian. Ví dụ: Dân số thành thị so với dân số nông thôn, dân số là nam so với dân số là nữ; giá trị tăng thêm của doanh nghiệp ngoài quốc doanh so với giá trị tăng thêm của doanh nghiệp quốc doanh; năng suất lúa của tỉnh X so với năng suất lúa của tỉnh Y; số học sinh đạt kết quả học tập khá giỏi so với số học sinh đạt kết quả trung bình,...

2.2.3. Số tương đối kế hoạch

Số tương đối kế hoạch là chỉ tiêu phản ánh mức cần đạt tới trong kỳ kế hoạch, hoặc mức đã đạt được so với kế hoạch được giao về một chỉ tiêu kinh tế - xã hội nào đó. Số tương đối kế hoạch được chia thành hai loại:

+ *Số tương đối nhiệm vụ kế hoạch*: Phản ánh quan hệ so sánh giữa mức độ đề ra trong kỳ kế hoạch với mức độ thực tế ở kỳ gốc của một chỉ tiêu kinh tế - xã hội.

+ *Số tương đối hoàn thành kế hoạch*: Phản ánh quan hệ so sánh giữa mức thực tế đã đạt được với mức kế hoạch trong kỳ về một chỉ tiêu kinh tế - xã hội.

2.2.4. Số tương đối kết cấu

Số tương đối kết cấu là chỉ tiêu phản ánh tỷ trọng của mỗi bộ phận chiếm trong tổng thể, tính được bằng cách đem so sánh mức độ tuyệt đối của từng bộ phận với mức độ tuyệt đối của toàn bộ tổng thể.

Số tương đối kết cấu thường được biểu hiện bằng số phần trăm. Ví dụ: Tỷ trọng của GDP theo từng ngành trong tổng GDP của nền kinh tế quốc dân; tỷ trọng dân số của từng giới nam hoặc nữ trong tổng số dân,...

2.2.5. Số tương đối cường độ

Số tương đối cường độ là chỉ tiêu biểu hiện trình độ phổ biến của một hiện tượng trong các điều kiện thời gian và không gian cụ thể.

Số tương đối cường độ tính được bằng cách so sánh mức độ của hai chỉ tiêu khác nhau nhưng có quan hệ với nhau. Số tương đối cường độ biểu hiện bằng đơn vị kép, do đơn vị tính ở tử số và ở mẫu số hợp thành. Số tương đối cường độ được tính toán và sử dụng rất phổ biến trong công tác thống kê. Các số tương đối trong số liệu thống kê thường gặp như mật độ dân số bằng tổng số dân (người) chia cho diện tích tự nhiên (km^2) với đơn vị tính là người / km^2 ; GDP bình quân đầu người bằng tổng GDP (nghìn đồng) chia cho dân số trung bình (người) với đơn vị tính là 1000đ/người; số bác sĩ tính bình quân cho một vạn dân bằng tổng số bác sĩ chia cho tổng số dân tính bằng vạn người với đơn vị tính là người /10000 người,...

2.3. SỐ BÌNH QUÂN (TRONG THỐNG KÊ)

Số bình quân là chỉ tiêu biểu hiện mức độ điển hình của một tổng thể gồm nhiều đơn vị cùng loại được xác định theo một tiêu thức nào đó. Số bình quân được sử dụng phổ biến trong thống kê để nêu lên đặc điểm chung nhất, phổ biến nhất của hiện tượng kinh tế - xã hội trong các điều kiện thời gian và không gian cụ thể. Ví dụ: Tiền lương bình quân một công nhân trong doanh nghiệp là mức lương phổ biến nhất, đại diện cho các mức lương khác nhau của công nhân trong doanh nghiệp; thu nhập bình quân đầu người của một địa bàn là mức thu nhập phổ biến nhất, đại diện cho các mức thu nhập khác nhau của mọi người trong địa bàn đó.

Số bình quân còn dùng để so sánh đặc điểm của những hiện tượng không có cùng một quy mô hay làm căn cứ để đánh giá trình độ đồng đều của các đơn vị tổng thể.

Xét theo vai trò đóng góp khác nhau của các thành phần tham gia bình quân hoá, số bình quân chung được chia thành số bình quân giản đơn và số bình quân gia quyền.

+ Số bình quân giản đơn: Được tính trên cơ sở các thành phần tham gia bình quân hoá có vai trò về qui mô (tần số) đóng góp như nhau.

+ Số bình quân gia quyền: Được tính trên cơ sở các thành phần tham gia bình quân hoá có vai trò về qui mô (tần số) đóng góp khác nhau.

Để tính được số bình quân chính xác và có ý nghĩa, điều kiện chủ yếu là nó phải được tính cho những đơn vị cùng chung một tính chất (thường gọi là tổng thể đồng chất). Muốn vậy, phải dựa trên cơ sở phân tổ thống kê một cách khoa học và chính xác. Đồng thời phải vận dụng kết hợp giữa số bình quân tổ với số bình quân chung.

Có nhiều loại số bình quân khác nhau. Trong thống kê kinh tế - xã hội thường dùng các loại sau: Số bình quân số học, số bình quân điều hoà, số bình quân hình học (số bình quân nhân), một và trung vị.

Dưới đây là từng loại số bình quân nêu trên.

2.3.1. Số bình quân số học

Số bình quân số học được tính bằng cách chia tổng các lượng biến (theo một tiêu thức nào đó) cho số đơn vị tổng thể.

Công thức:

a. Số bình quân số học giản đơn

$$\bar{x}_s = \frac{\sum_{i=1}^n x_i}{n} ; \quad (2.3.1a)$$

Trong đó: $\bar{x}_s$ - Số bình quân số học;

x_i ($i = 1, 2, \dots, n$) - Các trị số của lượng biến;

n - Số đơn vị tổng thể.

Ví dụ: Một tổ có 5 công nhân, năng suất lao động của từng công nhân từ 1 đến 5 như sau: 2000 nghìn đồng,

2500 nghìn đồng, 3000 nghìn đồng, 3000 nghìn đồng và 3500 nghìn đồng.

Năng suất bình quân của 5 công nhân là:

$$\bar{x}_s = \frac{2000 + 2500 + 3000 + 3000 + 3500}{5} = 2800 \text{ (nghìn đồng)}$$

b. Số bình quân số học gia quyền

$$\bar{x}_s = \frac{\sum_{i=1}^k x_i f_i}{\sum_{i=1}^k f_i} ; \quad (2.3.1b)$$

Trong đó: f_i - Quyền số của lượng biến x_i (số đơn vị tổng thể có lượng biến x_i) ($\sum_{i=1}^k f_i = n$).

Ví dụ: Một tổ học sinh có 10 học sinh, với kết quả học môn toán của các em như sau: Điểm 7 có 3 em; điểm 8 có 5 em và điểm 9 có 2 em. Vậy điểm môn toán bình quân của 10 em học sinh như sau:

$$\bar{x}_s = \frac{(7 \times 3) + (8 \times 5) + (9 \times 2)}{10} = 7,9 \text{ (điểm)}$$

2.3.2. Số bình quân điều hoà

Số bình quân tính được từ nghịch đảo của các lượng biến ($\frac{1}{x_i}$).

Công thức:

a. Số bình quân điều hoà giản đơn

$$\bar{x}_h = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}} ; \quad (2.3.2a)$$

Trong đó:

$\bar{x}_h$ - Số bình quân điều hoà;

x_i ($i=1,2,3,\dots,n$) - Các lượng biến;

n - Số đơn vị tổng thể (số lượng biến).

Ví dụ: Một tổ sản xuất có 5 công nhân ($n=5$) cùng sản xuất một loại sản phẩm và cùng làm việc trong một thời gian như nhau. Người công nhân thứ nhất sản xuất một sản phẩm hết 2 phút, người thứ hai sản xuất một sản phẩm hết 3 phút, người thứ ba sản xuất một sản phẩm hết 4 phút, người thứ tư sản xuất 1 sản phẩm hết 5 phút và người thứ năm sản xuất một sản phẩm hết 6 phút. Thời gian hao phí bình quân ($\bar{x}_h$) để sản xuất một sản phẩm của 5 công nhân bằng:

$$\bar{x}_h = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}} = \frac{5}{\frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6}} = 3,45 \text{ (phút/sản phẩm)}$$

b. Số bình quân điều hoà gia quyền

$$\bar{x}_h = \frac{\sum_{i=1}^k M_i}{\sum_{i=1}^k \frac{1}{M_i}} = \frac{\sum_{i=1}^k M_i}{\sum_{i=1}^k \frac{M_i}{x_i}} \quad ; \quad (2.3.2b)$$

Trong đó:

M_i - Quyền số ($M_i = x_i f_i$ với $i=1, 2, \dots, k$).

Ví dụ: Một phân xưởng sản xuất có 3 tổ công nhân. Tổ 1 sản xuất được 220 sản phẩm (M_1) và năng suất lao động mỗi công nhân là 11 sản phẩm (x_1); tổ 2 sản xuất được 264 sản phẩm (M_2) với năng suất lao động mỗi công nhân là 12 sản phẩm (x_2) và tổ 3 sản xuất được 312 sản phẩm (M_3) với năng suất lao động mỗi công nhân 13 sản phẩm (x_3).

Vậy năng suất lao động bình quân mỗi công nhân trong phân

xưởng là:

$$\bar{x}_h = \frac{\frac{220}{11} + \frac{264}{12} + \frac{312}{13}}{\frac{220}{11} + \frac{264}{12} + \frac{312}{13}} = \frac{796}{66} = 12,06 \text{ (sản phẩm)}$$

2.3.3. Số bình quân nhân

Số bình quân nhân tính được bằng cách khai căn bậc n của tích n lượng biến.

Công thức:

a. Số bình quân nhân giản đơn

$$\bar{x}_\Pi = \sqrt[n]{\prod_{i=1}^n x_i} \quad ; \quad (2.3.3a)$$

Trong đó:

$\bar{x}_\Pi$ - Số bình quân nhân;

x_i ($i=1, 2, \dots, n$) - Các lượng biến;

n - Số lượng biến;

Π - Ký hiệu của tích.

Ví dụ: Tốc độ phát triển sản xuất của tỉnh "X" từ năm 1998 đến năm 2002 như sau: 1,775; 1,289; 1,322; 1,307; 1,222.

Tốc độ phát triển bình quân năm ($\bar{x}_\Pi$) của tỉnh "X" từ năm 1998 đến 2002 là:

$$\bar{x}_\Pi = \sqrt[5]{1,775 \cdot 1,289 \cdot 1,322 \cdot 1,307 \cdot 1,222} = 1,367 \text{ hoặc } 136,7\%$$

b. Số bình quân nhân gia quyền

$$\bar{x}_\Pi = \sqrt[n]{\prod_{i=1}^k x_i^{f_i}} \quad ; \quad (2.3.3b)$$

Trong đó: f_i - Quyền số với $\sum_{i=1}^k f_i = n$.

Ví dụ: Trong thời gian 10 năm ($\sum_{i=1}^k f_i = 10$) tốc độ phát triển sản xuất của một tỉnh "X" như sau: 5 năm đầu, mỗi năm có tốc độ phát triển là 1,1; trong 3 năm tiếp theo, mỗi năm có tốc độ phát triển là 1,15; 2 năm cuối cùng, mỗi năm có tốc độ phát triển là 1,25. Vậy tốc độ phát triển bình quân ($\bar{x}_{\Pi}$) của tỉnh "X" mỗi năm thời kỳ 10 năm chính là số bình quân nhân gia quyền được tính như sau:

$$\bar{x}_{\Pi} = \sqrt[10]{(1,1)^5 \cdot (1,15)^3 \cdot (1,25)^2} = 1,144 \text{ hoặc } 114,4\%$$

Số bình quân nhân được áp dụng trong trường hợp các lượng biến có quan hệ tích số với nhau và thường được dùng để tính tốc độ phát triển bình quân trong thực tế công tác thống kê.

2.3.4. Một

Một là biểu hiện của một tiêu thức số lượng được gặp nhiều nhất trong một tổng thể hay trong một dãy số phân phối. Trong dãy số lượng biến xác định, một là lượng biến có tần số lớn nhất. Một dùng để biểu hiện mức độ phổ biến của hiện tượng. Ví dụ: Trong số lượng áo sơ mi các cỡ bán ra của một cửa hàng, số lượng áo cỡ 40 bán được nhiều nhất thì một chính là loại áo sơ mi cỡ 40. Một số ví dụ khác trong địa bàn điều tra về thu nhập của các hộ gia đình, số hộ có mức thu nhập 3 triệu đồng một tháng là nhiều nhất, thì mức thu nhập 3 triệu đồng chính là một; trong một doanh nghiệp số công nhân có mức năng suất lao động 5 triệu đồng một tháng là nhiều nhất, thì mức năng suất lao động 5 triệu đồng chính là một,...

Trong một dãy số lượng biến có khoảng cách tổ muốn tìm một, trước hết cần xác định tổ có một, tức là tổ có tần số lớn nhất, sau đó tính trị số gần đúng của một theo công thức sau:

$$M_0 = x_{M_0(\min)} + i_{M_0} \frac{f_{M_0} - f_{M_0-1}}{(f_{M_0} - f_{M_0-1}) + (f_{M_0} - f_{M_0+1})} \quad ; \quad (2.3.4a)$$

Trong đó:

- M_0 - Một;
- $x_{M_0(\min)}$ - Giới hạn dưới của tổ có một;
- i_{M_0} - Trị số khoảng cách tổ có một;
- f_{M_0-1} - Tần số của tổ đứng trước tổ có một;
- f_{M_0} - Tần số tổ có một;
- f_{M_0+1} - Tần số của tổ đứng sau tổ có một.

Ví dụ: Có tình hình về tiền lương bình quân một tháng của công nhân trong một doanh nghiệp như bảng 2.3.1:

Bảng 2.3.1: Lương của công nhân trong doanh nghiệp

Thứ tự tổ	Mức lương (1000 đ)	Số công nhân (Người)	Thứ tự tổ	Mức lương (1000 đ)	Số công nhân (Người)
A	1	2	A	1	2
1	400 - 500	20	4	700 - 800	160
2	500 - 600	60	5	800 - 900	60
3	600 - 700	90	6	900 - 1000	10

Từ số liệu bảng 2.3.1, ta thấy tổ thứ tư ($i = 4$) là tổ có một ($f_4 = 160$) và khi đó giới hạn dưới $x_{M_0(\min)} = 700$, khoảng cách của tổ có một: $i_{M_0} = 800 - 700 = 100$, tần số của tổ đứng trước tổ có một $f_{M_0-1} = 90$ và tần số của tổ đứng sau tổ có một $f_{M_0+1} = 60$. Áp dụng công thức 2.3.4a tính được một, hay mức lương phổ biến nhất của doanh nghiệp như sau:

$$M_0 = 700 + 100 \cdot \frac{160 - 90}{(160 - 90) + (160 - 60)} = 741,2 \text{ (nghìn đồng)}$$

Ghi chú: Trường hợp khoảng cách tổ không bằng nhau việc xác định một phải căn cứ vào mật độ phân phối.

Trong một dãy số lượng biến không có khoảng cách tổ thì một (M_0) là lượng biến có tần số lớn nhất.

Một biểu hiện mức độ phổ biến của hiện tượng, đồng thời bản thân nó không san bằng, bù trừ chênh lệch giữa các lượng biến, cho nên có thể dùng để thay thế số bình quân trong những trường hợp cần thiết, nhất là khi dãy số có những lượng biến quá lớn hoặc quá nhỏ. Tuy nhiên, như vậy một sẽ có nhược điểm là kém nhạy bén đối với sự biến thiên của mỗi tiêu thức.

Một chỉ vận dụng đối với tổng thể tương đối nhiều đơn vị, không nên vận dụng trong trường hợp phân phối có quá nhiều điểm tập trung hoặc không có điểm chính tập trung các đơn vị.

Một còn được dùng để khảo sát tính chất đều đặn của dãy số phân phối và chỉ tiêu đánh giá tính chất đều đặn của dãy số phân phối gọi là hệ số đối ứng (K_A), tính theo công thức:

$$K_A = \frac{\bar{x} - M_0}{\sigma}; \quad (2.3.4b)$$

Trong đó:

$\bar{x}$ - Số bình quân số học;

σ - Độ lệch tiêu chuẩn (nội dung và công thức tính độ lệch tiêu chuẩn sẽ được giải thích sau).

2.3.5. Số trung vị

Số trung vị là lượng biến của một tiêu thức nào đó đứng ở vị trí giữa trong dãy số lượng biến.

+ Nếu tổng thể có số quan sát là lẻ thì trung vị sẽ chính là trị số

của số quan sát ở vị trí chính giữa. Khi đó dãy số lượng biến được chia thành hai phần (phần trên và phần dưới số trung vị) và mỗi phần có số đơn vị tổng thể bằng nhau. Ví dụ: Tiền lương của 9 công nhân được sắp xếp theo thứ tự mức lương tăng dần: 500, 520, 550, 570, 580, 600, 630, 640, 650 (nghìn đồng) thì số trung vị chính là tiền lương của công nhân đứng ở vị trí thứ 5 (giữa của 9 người), tức là có mức lương 580 nghìn đồng.

+ Nếu tổng thể có số quan sát là chẵn thì trung vị sẽ là số bình quân giản đơn của 2 quan sát ở vị trí giữa. Ví dụ tiền lương của 12 công nhân được sắp xếp theo thứ tự mức lương tăng dần: 600, 610, 615, 630, 650, 655, 665, 680, 690, 695, 700, 720 (nghìn đồng) thì số trung vị sẽ là số bình quân giản đơn của 2 người đứng ở vị trí thứ 6 và thứ 7, tức là $(655 + 665) : 2 = 660$ (nghìn đồng).

Trong một dãy số lượng biến có khoảng cách tổ, muốn tìm số trung vị trước hết cần xác định tổ có số trung vị (tổ có chứa đơn vị đứng ở vị trí giữa). Sau đó tính trị số gần đúng của số trung vị theo công thức:

$$M_e = x_{Me(min)} + i_{Me} \frac{\frac{\sum f_i}{2} - S_{(Me-1)}}{f_{Me}}; \quad (2.3.5)$$

Trong đó:

M_e - Số trung vị;

$x_{Me(min)}$ - Giới hạn dưới của tổ có số trung vị;

i_{Me} - Trị số của khoảng cách tổ của tổ có số trung vị;

$\sum f_i$ - Tổng các tần số (Số đơn vị tổng thể) trong dãy số;

$S_{(Me-1)}$ - Tổng các tần số của các tổ đứng trước tổ có số trung vị;

f_{Me} - Tần số của tổ có số trung vị.

Trở lại ví dụ trên (xem số liệu bảng 2.3.1) ta thấy tổ có chứa đơn

vị đứng giữa là tổ 4 ($i = 4$) và khi đó giới hạn dưới của tổ có số trung vị: $x_{Me(min)} = 700$, trị số khoảng cách tổ của tổ có trung vị: $i_{Me} = 800 - 700 = 100$, tổng các tần số trong dãy số $\sum f_i = 400$, tổng các tần số của các tổ đứng trước tổ có trung vị: $S_{(Me-1)} = 170$, tần số của tổ có trung vị: $f_{Me} = 160$. Áp dụng công thức 2.3.5 ta tính được số trung vị:

$$M_e = 700 + 100 \frac{\frac{400}{2} - 170}{160} = 718,8 \text{ (nghìn đồng)}$$

Số trung vị có thể dùng để bổ sung hoặc thay thế cho số bình quân số học khi không biết chính xác toàn bộ các lượng biến; chỉ cần đảm bảo được sự phân phối của các đơn vị theo thứ tự tăng dần của lượng biến là có thể tính được số trung vị.

2.4. ĐỘ BIẾN THIÊN CỦA TIÊU THỨC

Độ biến thiên của tiêu thức dùng để đánh giá mức độ đại diện của số bình quân đối với tổng thể được nghiên cứu. Trị số này tính ra càng lớn, độ biến thiên của tiêu thức càng lớn do đó mức độ đại diện của số bình quân đối với tổng thể càng thấp và ngược lại.

Quan sát độ biến thiên tiêu thức trong dãy số lượng biến sẽ thấy nhiều đặc trưng về phân phối, kết cấu, tính đồng đều của tổng thể.

Độ biến thiên của tiêu thức được sử dụng nhiều trong nghiên cứu thống kê như phân tích biến thiên cũng như mối liên hệ của hiện tượng, dự đoán thống kê, điều tra chọn mẫu,...

Khi nghiên cứu độ biến thiên của tiêu thức, thống kê thường dùng các chỉ tiêu như khoảng biến thiên, độ lệch tuyệt đối bình quân, phương sai, độ lệch tiêu chuẩn và hệ số biến thiên. Dưới đây là nội dung và phương pháp tính của các chỉ tiêu đó.

2.4.1. Khoảng biến thiên

Khoảng biến thiên (còn gọi là toàn cự) là chỉ tiêu được tính bằng hiệu số giữa lượng biến lớn nhất và lượng biến nhỏ nhất của một dãy số lượng biến. Khoảng biến thiên càng lớn, mức độ biến động của chỉ tiêu càng lớn. Ngược lại, khoảng biến thiên nhỏ, mức độ biến động của chỉ tiêu thấp, tức là mức độ đồng đều của chỉ tiêu cao.

Công thức:

$$R = X_{\max} - X_{\min} \quad ; \quad (2.4.1)$$

Trong đó:

R - Toàn cự;

$X_{\max}$ - Lượng biến có trị số lớn nhất;

$X_{\min}$ - Lượng biến có trị số nhỏ nhất.

Ví dụ: Thu nhập của hộ gia đình như bảng 2.4.1:

Bảng 2.4.1: Thu nhập của hộ gia đình

Hộ	1	2	3	4	5	6	7	8
Thu nhập (1000 đồng)	6000	7000	85000	86000	9000	9100	9500	10000

Từ số liệu bảng 2.4.1 sử dụng công thức 2.4.1 ta tính được khoảng biến thiên:

$$R = 10000 - 6000 = 4000 \text{ (nghìn đồng)}$$

Khoảng biến thiên phản ánh khoảng cách biến động của tiêu thức tuy tính toán đơn giản song phụ thuộc vào lượng biến lớn nhất và nhỏ nhất của tiêu thức, tức là không tính gì đến mức độ khác nhau của các lượng biến còn lại trong dãy số.

2.4.2. Độ lệch tuyệt đối bình quân

Độ lệch tuyệt đối bình quân là số bình quân số học của các độ

lệch tuyệt đối giữa các lượng biến với số bình quân số học của các lượng biến đó.

Công thức:

$$\text{Trường hợp tính giản đơn} \quad \bar{d} = \frac{\sum |x_i - \bar{x}|}{n} \quad ; \quad (2.4.2a)$$

$$\text{Trường hợp có quyền số} \quad \bar{d} = \frac{\sum |x_i - \bar{x}| f_i}{\sum f_i} \quad ; \quad (2.4.2b)$$

Trong đó:

$\bar{d}$ - Độ lệch tuyệt đối bình quân;

x_i ($i = 1, 2, 3, \dots, n$ nếu tính giản đơn; $i = 1, 2, \dots, k$ nếu tính gia quyền) - Các trị số của lượng biến;

$\bar{x}$ - Số bình quân số học;

f_i - Quyền số của từng lượng biến x_i ;

n - Tổng số lượng biến ($n = \sum_{i=1}^k f_i$).

Chỉ tiêu này biểu hiện độ biến thiên của tiêu thức nghiên cứu một cách đầy đủ hơn khoảng biến thiên. Qua đó phản ánh rõ nét hơn tính chất đồng đều của tổng thể: vì nó tính đến độ lệch của tất cả các lượng biến. Về cách tính cũng tương đối đơn giản, nhưng có đặc điểm là phải lấy giá trị tuyệt đối (giá trị dương) của chênh lệch.

Ví dụ: Có số liệu về năng suất lao động năm của công nhân trong một doanh nghiệp như bảng 2.4.2:

Bảng 2.4.2: Năng suất lao động của công nhân trong doanh nghiệp

STT	Năng suất lao động năm (Triệu đồng)	Số công nhân (Nghìn người)	STT	Năng suất lao động năm (Triệu đồng)	Số công nhân (Nghìn người)
-----	-------------------------------------	----------------------------	-----	-------------------------------------	----------------------------

	/người)			/người)	người)
A	1	2	A	1	2
1	10	10	4	25	10
2	15	20	5	35	10
3	20	50			

Từ số liệu bảng 2.4.2 sử dụng công thức 2.3.1b và 2.4.2b ta tính được:

a. Số bình quân

$$\bar{x} = \frac{(10.10) + (15.20) + (20.50) + (25.10) + (35.10)}{10 + 20 + 50 + 10 + 10} = 20$$

b. Độ lệch tuyệt đối bình quân

$$\begin{aligned} \bar{d} &= \frac{|10 - 20|10 + |15 - 20|20 + |20 - 20|50 + |25 - 20|10 + |35 - 20|10}{10 + 20 + 50 + 10 + 10} \\ &= \frac{400}{100} = 4 \end{aligned}$$

2.4.3. Phương sai

Phương sai là số bình quân số học của bình phương các độ lệch giữa các lượng biến với số bình quân số học của các lượng biến đó.

Công thức:

$$\text{Trường hợp tính giản đơn} \quad \sigma^2 = \frac{\sum (x_i - \bar{x})^2}{n} \quad ; \quad (2.4.3a)$$

$$\text{Trường hợp có quyền số} \quad \sigma^2 = \frac{\sum (x_i - \bar{x})^2 f_i}{\sum f_i} \quad ; \quad (2.4.3b)$$

Trong đó:

σ^2 - Phương sai;

x_i - ($i=1,2,..., n$ trường hợp giản đơn và $i = 1,2,..., k$ trường hợp có quyền số) - Các trị số của lượng biến;

$\bar{x}$ - Số bình quân số học;

f_i - Quyền số của từng lượng biến x_i ;

n - Tổng số lượng biến ($n = \sum f_i$)

Cũng từ số liệu về năng suất lao động của công nhân một doanh nghiệp trong bảng 2.4.2 ở trên áp dụng công thức 2.4.3b ta tính được phương sai (trường hợp có quyền số):

$$\sigma^2 = \frac{(10-20)^2.10+(15-20)^2.20+(20-20)^2.50+(25-20)^2.10+(35-20)^2.10}{10+20+50+10+10}$$

$$= \frac{1000+500+250+2250}{100} = \frac{4000}{100} = 40$$

2.4.4. Độ lệch chuẩn

Độ lệch chuẩn là căn bậc 2 của phương sai cho biết bình quân giá trị của các lượng biến cách giá trị trung bình chung là bao nhiêu đơn vị.

Công thức tính:

Trường hợp giản đơn $\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n}}$; (2.4.4a)

Trường hợp có quyền số

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum (x_i - \bar{x})^2 f_i}{\sum f_i}}$$
 ; (2.4.4b)

Theo ví dụ về phương sai tính được ở trên: ($\sigma^2 = 40$) thì độ lệch chuẩn $\sigma = \sqrt{40} = 6,32$ (triệu đồng).

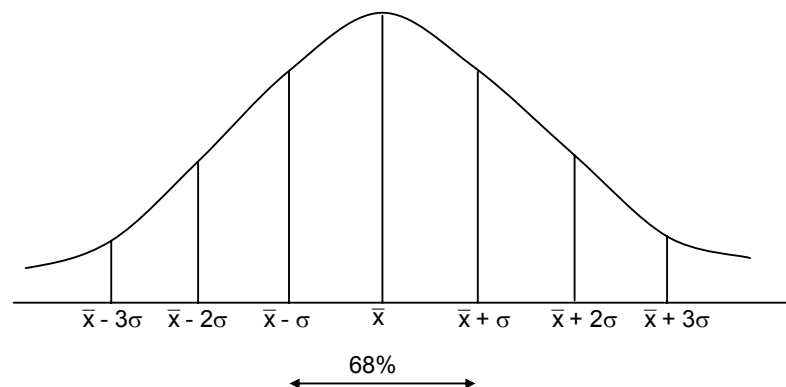
Độ lệch chuẩn cho phép ta xác định vị trí phân bố của dãy số

trong mối quan hệ với số trung bình. Theo định lý của Chebyshev:

- Bất kỳ sự sắp xếp nào cũng có ít nhất 75% giá trị sẽ rơi vào trong khoảng cộng trừ hai lần độ lệch chuẩn ($\pm 2\sigma$) từ số trung bình và có ít nhất 89% giá trị sẽ nằm trong khoảng cộng trừ 3 lần độ lệch chuẩn ($\pm 3\sigma$) từ số trung bình.

- Đối với phân bố chuẩn sẽ có khoảng 68% giá trị của tổng thể chung rơi vào trong khoảng tin cậy độ lệch chuẩn ($\pm \sigma$) từ số trung bình, 95% giá trị sẽ rơi vào trong khoảng tin cậy hai lần độ lệch chuẩn ($\pm 2\sigma$) từ số trung bình và 99% giá trị nằm trong khoảng tin cậy ba lần độ lệch chuẩn ($\pm 3\sigma$) từ số trung bình (xem hình vẽ 2.4.1).

Hình 2.4.1: Đường biểu diễn phân phối chuẩn



Độ lệch chuẩn là một trong những chỉ tiêu thường dùng nhất để biểu hiện độ biến thiên của tổng thể được nghiên cứu và đánh giá trình độ đồng đều của tổng thể được nghiên cứu.

Độ lệch chuẩn có nhiều ứng dụng quan trọng trong các quá trình tính toán và phân tích thống kê như: Xác định số mẫu cần chọn trong điều tra chọn mẫu, tính hệ số tương quan hoặc tỷ số tương quan, tính hệ số biến thiên, v.v...

Vì độ lệch chuẩn là căn bậc 2 của phương sai, nên khi nói đến vai

trò của độ lệch chuẩn thì cũng chính là nói đến vai trò của phương sai. Hay nói cách khác muốn có độ lệch chuẩn nhất thiết phải có phương sai.

2.4.5. Hệ số biến thiên

Hệ số biến thiên là chỉ tiêu tương đối phản ánh mối quan hệ so sánh giữa độ lệch chuẩn với số bình quân số học.

Công thức:

$$V = \frac{\sigma}{\bar{x}} \quad ; \quad (2.4.5a)$$

Trong đó:

V - Hệ số biến thiên;

σ - Độ lệch chuẩn;

$\bar{x}$ - Số bình quân số học.

Ví dụ: Khi độ lệch chuẩn $\sigma = 6,32$; số bình quân số học $\bar{x} = 20$ thì sẽ có hệ số biến thiên là: $V = \frac{6,32}{20} = 0,316$ hoặc 31,6%.

Hệ số biến thiên cũng dùng để đánh giá độ biến thiên của tiêu thức và tính chất đồng đều của tổng thể. Hệ số này biểu hiện bằng số tương đối nên còn có thể được dùng để so sánh cả những chỉ tiêu cùng loại nhưng ở các quy mô khác nhau như so sánh độ đồng đều về thu nhập bình quân của hộ gia đình ở một tỉnh miền núi (có thu nhập thấp và số hộ ít hơn) với thu nhập bình quân của hộ gia đình ở thủ đô Hà Nội (có mức thu nhập cao hơn và số hộ nhiều hơn), đặc biệt để so sánh được những chỉ tiêu của các hiện tượng khác nhau và có đơn vị đo lường khác nhau như so sánh hệ số biến thiên về bậc thợ với hệ số biến thiên về tiền lương bình quân, hệ số biến thiên về năng suất lao động bình quân, so sánh hệ số biến thiên về chỉ tiêu thu nhập của hộ gia đình với hệ số biến thiên về chi tiêu của hộ gia đình,...

Hệ số biến thiên còn có thể tính theo độ lệch tuyệt đối bình quân,

nhưng hệ số biến thiên tính theo độ lệch chuẩn thường được sử dụng rộng rãi hơn, tuy phần tính toán có phức tạp hơn phải sử dụng MTĐT.

Hệ số biến thiên tính theo độ lệch tuyệt đối bình quân có công thức tính:

$$V = \frac{\bar{d}}{\bar{x}} \quad ; \quad (2.4.5b)$$

Trong đó: $\bar{d}$ - Độ lệch tuyệt đối bình quân.

2.5. MỨC ĐỒNG ĐỀU CỦA PHÂN PHỐI

Để xác định mức độ biến thiên đồng đều hoặc bất bình đẳng của phân phối có thể dùng nhiều phương pháp, nhưng trong thống kê thường sử dụng đường cong Lorenz và hệ số GINI.

2.5.1. Đường cong Lorenz

Đó là một loại đồ thị dùng để biểu diễn mức độ thiếu đồng đều hoặc bất bình đẳng của phân phối. Ví dụ, nghiên cứu phân phối thu nhập của dân cư, đường cong Lorenz biểu thị quan hệ giữa tỷ lệ phần trăm số dân cư và tỷ lệ phần trăm thu nhập của các nhóm dân cư đó. Nghiên cứu phân bố về dân số, đường cong Lorenz biểu thị quan hệ giữa phần trăm diện tích tự nhiên của từng địa phương với phần trăm của dân số của các địa phương đó. Khi nghiên cứu phân phối thu nhập của dân cư, trên đồ thị, trục hoành biểu thị tỷ lệ phần trăm cộng dồn của số dân cư từ 0% đến 100% được sắp xếp theo thứ tự nhóm dân cư có thu nhập tăng dần và trục tung biểu thị tỷ lệ phần trăm cộng dồn thu nhập của các nhóm dân cư từ 0% đến 100%.

Vì các nhóm dân cư được sắp xếp theo thứ tự từ nhóm có thu nhập thấp nhất đến nhóm có thu nhập cao nhất nên tỷ lệ phần trăm cộng dồn số dân của các nhóm dân cư luôn luôn lớn hơn phần trăm cộng dồn thu nhập tương ứng của nhóm, do vậy đường cong Lorenz

luôn nằm dưới đường nghiêng 45^0 và có mặt lõm hướng lên trên (xem hình vẽ theo ví dụ). Đường cong Lorenz càng lõm (diện tích hình A càng lớn) thì sự bất bình đẳng càng cao và ngược lại. Nếu tất cả các nhóm dân cư có mức thu nhập giống nhau, khi đó đường cong Lorenz sẽ trùng với đường nghiêng 45^0 và được gọi là đường bình đẳng tuyệt đối.

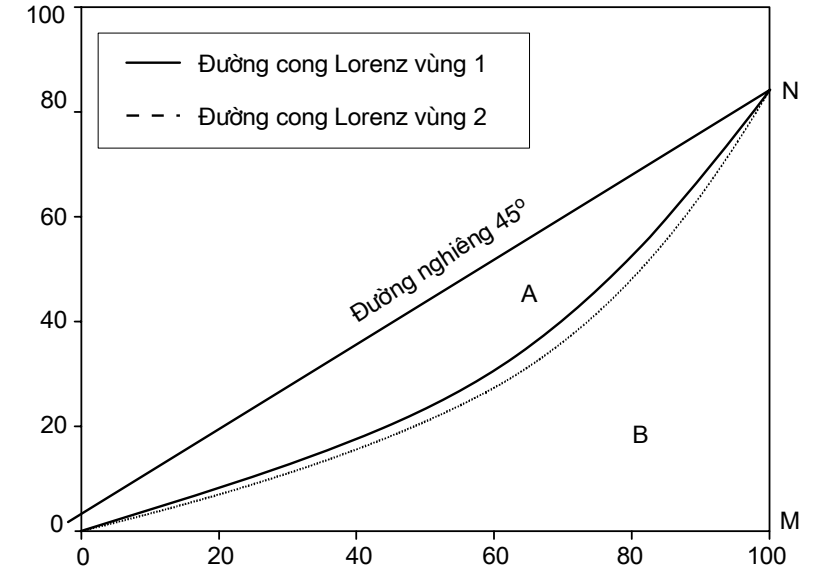
Ví dụ: Có số liệu về thu nhập của các tầng lớp dân cư của 2 vùng nước ta trong cùng một thời kỳ như bảng 2.5.1:

Bảng 2.5.1: Thu nhập của dân cư trong 2 vùng

Phần trăm dân số theo mức giàu, nghèo	Phần trăm thu nhập		Phần trăm cộng dồn của dân số	Phần trăm cộng dồn của thu nhập	
	Vùng 1	Vùng 2		Vùng 1	Vùng 2
20% nghèo nhất	7	6	20	7	6
20% dưới trung bình	12	10	40	19	16
20% trung bình	18	17	60	37	33
20% khá	25	26	80	62	59
20% giàu	38	41	100	100	100

Biểu diễn mức độ chênh lệch về thu nhập của 2 vùng trên cùng một hệ tọa độ như sơ đồ 2.5.1:

Sơ đồ 2.5.1: Đường cong Lorenz của hai vùng



Hai đường cong trên cho ta một nhận biết về sự bất bình đẳng theo thu nhập của dân cư: Vùng 1 có mức độ chênh lệch nhỏ hơn vùng 2 vì khoảng cách từ đường nghiêng 45^0 tới đường cong Lorenz 1 gần hơn khoảng cách tới đường cong Lorenz 2.

Đường cong Lorenz không chỉ giúp ta so sánh sự biến động giữa các vùng mà còn giúp ta so sánh sự biến động theo thời gian. Muốn vậy, người ta vẽ các đường cong Lorenz của các năm khác nhau trong cùng một vùng trên cùng một hệ trục tọa độ.

2.5.2. Hệ số GINI

Hệ số GINI là số đo về sự bất bình đẳng của phân phối (thường là phân phối thu nhập của dân cư), được biểu hiện bằng tỷ lệ so sánh giữa phần diện tích giới hạn bởi đường nghiêng 45^0 và đường cong Lorenz với toàn bộ diện tích tam giác OMN. Nếu gọi A là phần diện tích giới hạn bởi đường nghiêng 45^0 (ON) với đường cong Lorenz và

B là diện tích còn lại của tam giác OMN thì ta có hệ số GINI (G):

$$G = \frac{A}{A + B} \quad ; \quad (2.5.1a)$$

Nếu đường cong Lorenz trùng với đường thẳng 45° (đường bình đẳng tuyệt đối) thì hệ số GINI bằng 0 (vì $A = 0$), xã hội có sự phân phối bình đẳng tuyệt đối. Nếu đường cong Lorenz trùng với trục hoành, hệ số GINI bằng 1 (vì $B = 0$), xã hội có sự phân phối bất bình đẳng tuyệt đối. Như vậy $0 \leq G \leq 1$.

Khi nghiên cứu về sự bất bình đẳng về thu nhập của dân cư, khi có số liệu về thu nhập và số người tương ứng chia theo các nhóm dân cư có mức thu nhập khác nhau, công thức tính hệ số GINI như sau:

$$G = 1 - \frac{\sum_{i=1}^n P_i (Q_i + Q_{i-1})}{100000} \quad ; \quad (2.5.1b)$$

Trong đó:

P_i - Tỷ lệ số người của nhóm dân thứ i

Q_i và Q_{i-1} - Tỷ lệ cộng dồn thu nhập đến nhóm dân cư thứ i và $i - 1$

Giả sử có số liệu về thu nhập của các nhóm dân cư một vùng trong năm như bảng 2.5.2.

Bảng 2.5.2: Bảng tính hệ số GINI

Thứ tự nhóm (i)	TNBQ 1 người (1000đ)	Tỷ lệ số người của từng nhóm ($P_i - \%$)	Tỷ lệ thu nhập của từng nhóm ($Q_i - \%$)	Tỷ lệ cộng dồn (%)		$Q+Q_{-1}$	$P_i(Q+Q_{-1})$
				Dân số (P)	Thu nhập (Q)		
A	1	2	3	4	5	6	7=2.6
1	550	20	11,46	20,00	11,46	11,46	229
2	650	18	13,54	38,00	25,00	36,46	656
3	750	20	15,63	58,00	40,63	65,63	1.313
4	850	16	17,71	74,00	58,33	98,96	1.583
5	950	15	19,79	89,00	78,13	136,46	2.047
6	1050	11	21,88	100,00	100,00	178,13	1.959
Tổng	4800	100	100	x	x	x	7.788

Thay số liệu vào công thức 2.5.2 ta tính được:

$$G = 1 - \frac{7788}{100000} = 1 - 0,7788 = 0,2213$$

Nếu như đường cong Lorenz giúp ta nhận biết bằng trực giác về tính chất và sự khác nhau về bất bình đẳng trong phân phối, thì hệ số GINI cho phép ta xác định mức độ bất bình đẳng đó đến đâu, với con số cụ thể là bao nhiêu.

Hệ số GINI là một số không âm ($0 \leq G \leq 1$); hệ số này càng nhỏ thì sự bình đẳng trong phân phối càng lớn và ngược lại hệ số này càng lớn thì sự bình đẳng trong phân phối càng nhỏ.

PHẦN BA

MỘT SỐ PHƯƠNG PHÁP THƯỜNG DÙNG TRONG PHÂN TÍCH THỐNG KÊ

Phân tích thống kê là giai đoạn cuối cùng của quá trình nghiên cứu thống kê, từ các biểu hiện về lượng nhằm nêu lên một cách tổng hợp bản chất và tính quy luật của các hiện tượng và quá trình kinh tế - xã hội trong các điều kiện thời gian và không gian cụ thể. Khi phân tích thống kê, người ta căn cứ vào các tài liệu báo cáo và điều tra đã được tổng hợp để tính các chỉ tiêu cần thiết, so sánh và biểu hiện các chỉ tiêu đó dưới dạng bảng số liệu hoặc đồ thị thống kê nhờ vào sự hỗ trợ của các phương pháp chuyên môn của khoa học thống kê, rút ra những kết luận đáp ứng mục đích nghiên cứu và đề xuất các biện pháp giải quyết.

Trong thống kê kinh tế - xã hội, nhiệm vụ chủ yếu của phân tích là đánh giá tình hình thực hiện các mục tiêu, chỉ ra những nguyên nhân hoàn thành hoặc không hoàn thành các mục tiêu, nêu rõ sự biến động và xu hướng phát triển của hiện tượng nghiên cứu trong mối quan hệ với các hiện tượng có liên quan, phát hiện ra các năng lực tiềm tàng có thể khai thác trong nền kinh tế, chỉ ra những mặt cân đối lớn, những mặt thuận lợi và khó khăn, những yếu tố thúc đẩy hoặc kìm hãm sự phát triển kinh tế - xã hội,...

Trong phân tích thống kê, không có mẫu báo cáo phân tích nào có thể áp dụng cho mọi trường hợp; mà tùy thuộc vào mục đích nghiên cứu, vào điều kiện cụ thể về nội dung và đặc điểm của hiện tượng, về nguồn số liệu hiện có mà xây dựng những mô hình phân tích phù hợp trên cơ sở áp dụng một cách linh hoạt các phương pháp phân tích thống kê. Trong đó các phương pháp thường được sử dụng là: Phương pháp phân tổ, phương pháp đồ thị, phương pháp dãy số biến động theo thời gian, phương pháp hồi quy tương quan, phương pháp chỉ số và

phương pháp cân đối.

3.1. PHƯƠNG PHÁP PHÂN TỔ THỐNG KÊ

3.1.1. Khái niệm phân tổ thống kê và tiêu thức phân tổ

Phân tổ thống kê là căn cứ vào một (hay một số) tiêu thức nào đó để phân chia tổng thể thống kê thành các tổ (tiểu tổ) có tính chất khác nhau. Ví dụ, phân chia nhân khẩu trong nước thành các tổ nam và nữ (căn cứ vào giới tính), thành các tổ có độ tuổi khác nhau (căn cứ vào độ tuổi), v.v... Một ví dụ khác: Phân chia chỉ tiêu giá trị tăng thêm của sản xuất công nghiệp thành các tổ là kinh tế nhà nước và kinh tế ngoài nhà nước (căn cứ vào hình thức sở hữu), thành các ngành công nghiệp riêng biệt (căn cứ vào hoạt động sản xuất công nghiệp), v.v...

Phân tổ thống kê là phương pháp cơ bản của tổng hợp thống kê, là một trong những phương pháp quan trọng của phân tích thống kê, đồng thời là cơ sở để vận dụng các phương pháp phân tích thống kê khác như phương pháp chỉ số, phương pháp tương quan, phương pháp cân đối,...

Tiêu thức thống kê (đặc điểm của đơn vị tổng thể để nhận thức hiện tượng nghiên cứu) được chọn làm căn cứ để phân tổ thống kê gọi là tiêu thức phân tổ. Tiêu thức phân tổ thống kê được chia thành 2 loại: Tiêu thức số lượng và tiêu thức thuộc tính.

Tiêu thức số lượng là tiêu thức có thể biểu diễn được bằng con số, ví dụ độ tuổi, thu nhập bình quân của hộ gia đình, trình độ văn hoá, mức năng suất lao động, tiền lương bình quân,...

Tiêu thức thuộc tính là tiêu thức không thể biểu hiện được bằng con số, ví dụ giới tính, nghề nghiệp, dân tộc, tôn giáo,...

3.1.2. Các loại phân tổ và cách thức tiến hành phân tổ

Trong thống kê, có thể phân tổ theo một tiêu thức (gọi là phân tổ đơn) hoặc phân tổ theo hai hay nhiều tiêu thức (gọi là phân tổ kết

hợp).

a. Phân tổ theo một tiêu thức

Phân tổ theo một tiêu thức là cách phân tổ đơn giản nhất và cũng thường được sử dụng nhất.

Cách tiến hành phân tổ, thường theo các bước sau:

+ *Chọn tiêu thức phân tổ:*

Chọn tiêu thức để phân tổ là vấn đề mang tính cốt lõi của phân tổ thống kê, vì phân tổ theo các tiêu thức khác nhau sẽ đáp ứng những mục đích nghiên cứu khác nhau, biểu hiện các khía cạnh khác nhau của tập hợp thông tin. Phải căn cứ vào mục đích nghiên cứu và bản chất của hiện tượng để xác định tiêu thức phân tổ cho phù hợp, đồng thời cần phải xét đến điều kiện cụ thể của hiện tượng.

+ *Xác định số tổ và khoảng cách tổ:*

Số lượng tổ phụ thuộc vào số lượng thông tin và phạm vi biến động của tiêu thức nghiên cứu. Lượng thông tin càng nhiều, phạm vi biến động của tiêu thức càng lớn thì càng phải phân làm nhiều tổ.

- Phân tổ theo tiêu thức thuộc tính.

Ở đây sự khác nhau giữa các tổ được biểu hiện bằng sự khác nhau giữa các loại hình. Nếu các loại hình tương đối ít, ta có thể coi mỗi loại hình là một tổ, tức là có bao nhiêu loại hình sẽ có bấy nhiêu tổ. Trường hợp số loại hình thực tế có nhiều, nếu như coi mỗi loại hình là một tổ thì số tổ sẽ quá nhiều, không thể khái quát chung được, cũng như không nêu được đặc điểm khác nhau giữa các tổ, cho nên cần phải ghép những loại hình giống nhau hoặc gần giống nhau vào cùng một tổ.

- Phân tổ theo tiêu thức số lượng.

Phân tổ theo tiêu thức số lượng là phân các đơn vị của tổng thể có lượng biến tương ứng với trị số khác nhau của tiêu thức phân tổ vào các tổ khác nhau.

Trường hợp sự biến thiên về lượng giữa các đơn vị không chênh lệch nhau nhiều và lượng biến thiên của tiêu thức phân tổ chỉ thay đổi trong phạm vi hẹp và biến động rời rạc như số lượng người trong gia đình, số điểm kết quả học tập của học sinh, số máy do công nhân phụ trách, v.v... thì có thể mỗi lượng biến là cơ sở để hình thành một tổ, hoặc ghép một số lượng biến vào một tổ tùy theo đặc tính của hiện tượng và mục đích nghiên cứu. *Ví dụ:* Phân tổ học sinh theo điểm kết quả học tập, ta có thể phân thành 10 tổ hoặc phân thành 5 tổ: Yếu, kém, trung bình, khá và giỏi.

Trường hợp lượng biến của tiêu thức biến thiên lớn, nếu mỗi lượng biến hình thành một tổ thì số tổ sẽ quá nhiều, đồng thời không nói rõ sự khác nhau về chất giữa các tổ. Trong trường hợp này cần chú ý tới mối liên hệ giữa lượng và chất trong phân tổ. Nghĩa là phải xem sự thay đổi về lượng đến mức độ nào thì bản chất của hiện tượng mới thay đổi và làm nảy sinh ra tổ khác. Như vậy mỗi tổ sẽ bao gồm một phạm vi lượng biến, có hai giới hạn: Giới hạn dưới là lượng biến nhỏ nhất và giới hạn trên là lượng biến lớn nhất của tổ; nếu vượt quá giới hạn này thì chất lượng thay đổi và chuyển sang tổ khác. Trị số chênh lệch giữa giới hạn trên và giới hạn dưới của mỗi tổ gọi là khoảng cách tổ (khoảng cách tổ có thể bằng nhau hoặc không bằng nhau).

Việc xác định khoảng cách tổ đều nhau hay không đều nhau là phải căn cứ vào đặc điểm của hiện tượng nghiên cứu. Phân tổ phải đảm bảo các đơn vị phân phối vào một tổ đều có cùng một tính chất và sự khác nhau về lượng giữa các tổ phải nêu rõ sự khác nhau về chất giữa các tổ. Trong thực tế, sự thay đổi về lượng của các bộ phận trong hiện tượng thường không diễn ra một cách đều đặn. Do đó trong rất nhiều trường hợp nghiên cứu phải phân tổ theo khoảng cách tổ không đều nhau. Riêng đối với các hiện tượng tương đối đồng nhất và lượng biến trên các đơn vị thay đổi một cách đều đặn, thì thường phân tổ với khoảng cách tổ đều nhau. Cách phân tổ này tạo điều kiện thuận lợi cho việc vận dụng các công thức toán học và dễ dàng trình bày số liệu trên các đồ thị thống kê. Việc phân tổ với khoảng cách tổ đều nhau tương

đối đơn giản và trị số khoảng cách tổ được xác định như sau:

$$\text{Khoảng cách tổ} = \frac{\text{Lượng biến lớn nhất} - \text{Lượng biến nhỏ nhất}}{\text{Số tổ cần thiết}}$$

+ *Phân các đơn vị vào các tổ tương ứng:*

Căn cứ vào lượng biến của từng đơn vị để phân đơn vị đó vào tổ có trị số của tiêu thức theo khoảng cách tổ phù hợp đã được xác định ở trên.

+ *Xác định tần số phân phối:*

Trên cơ sở số liệu đã phân tổ dễ dàng xác định được số đơn vị (tần số) của từng tổ. Hiện nay máy tính có thể giúp ta xác định các đại lượng trong phân tổ một cách rất thuận tiện và nhanh chóng.

b. Phân tổ theo nhiều tiêu thức

Phân tổ theo nhiều tiêu thức (còn gọi là phân tổ kết hợp) cũng được tiến hành giống như phân tổ theo một tiêu thức. Trước tiên phải xác định cần phân tổ theo những tiêu thức nào. Muốn chọn tiêu thức phân tổ phù hợp phải căn cứ vào mục đích nghiên cứu, vào bản chất của hiện tượng, vào mối liên hệ giữa các tiêu thức... Sau đó tiếp tục xác định xem tiêu thức nào phân trước, tiêu thức nào phân sau và theo mỗi tiêu thức sẽ phân làm bao nhiêu tổ.

Có thể phân tổ theo 2, 3, 4 tiêu thức hoặc nhiều hơn nữa. Song khi phân tổ phải căn cứ vào mục đích nghiên cứu và điều kiện số liệu để chọn bao nhiêu tiêu thức phân tổ cho phù hợp và chọn những tiêu thức nào cho có ý nghĩa nhất.

Trong thực tế công tác thống kê phân tổ theo hai hoặc ba tiêu thức là thường gặp nhất; ví dụ dân số phân theo độ tuổi và giới tính, GDP phân theo khu vực và ngành kinh tế,... (2 tiêu thức); cán bộ khoa học công nghệ phân theo trình độ chuyên

môn, giới tính và lĩnh vực hoạt động khoa học; khách du lịch phân theo quốc tịch, mục đích du lịch và giới tính,... (theo 3 tiêu thức).

3.2. PHƯƠNG PHÁP ĐỒ THỊ THỐNG KÊ

Phương pháp đồ thị thống kê là phương pháp trình bày và phân tích các thông tin thống kê bằng các biểu đồ, đồ thị và bản đồ thống kê. Phương pháp đồ thị thống kê sử dụng con số kết hợp với các hình vẽ, đường nét và màu sắc để trình bày các đặc điểm số lượng của hiện tượng. Chính vì vậy, ngoài tác dụng phân tích giúp ta nhận thức được những đặc điểm cơ bản của hiện tượng bằng trực quan một cách dễ dàng và nhanh chóng, đồ thị thống kê còn là một phương pháp trình bày các thông tin thống kê một cách khái quát và sinh động, chứa đựng tính mỹ thuật; thu hút sự chú ý của người đọc, giúp người xem dễ hiểu, dễ nhớ nên có tác dụng tuyên truyền cổ động rất tốt. Đồ thị thống kê có thể biểu thị:

- Kết cấu của hiện tượng theo tiêu thức nào đó và sự biến đổi của kết cấu.
- Sự phát triển của hiện tượng theo thời gian.
- So sánh các mức độ của hiện tượng.
- Mối liên hệ giữa các hiện tượng.
- Trình độ phổ biến của hiện tượng.
- Tình hình thực hiện kế hoạch.

Trong công tác thống kê thường dùng các loại đồ thị: Biểu đồ hình cột, biểu đồ tượng hình, biểu đồ diện tích (hình vuông, hình tròn, hình chữ nhật), đồ thị đường gấp khúc và biểu đồ hình màng nhện.

3.2.1. Biểu đồ hình cột

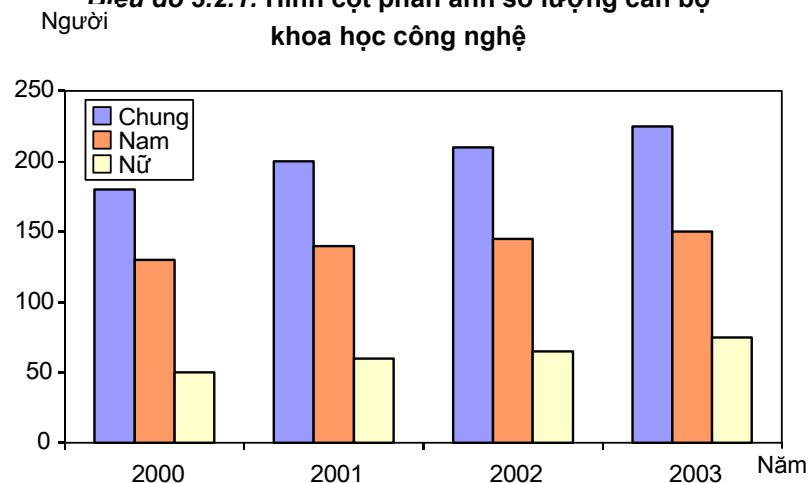
Biểu đồ hình cột là loại biểu đồ biểu hiện các tài liệu thống kê

bằng các hình chữ nhật hay khối chữ nhật thẳng đứng hoặc nằm ngang có chiều rộng và chiều sâu bằng nhau, còn chiều cao tương ứng với các đại lượng cần biểu hiện.

Biểu đồ hình cột được dùng để biểu hiện quá trình phát triển, phản ánh cơ cấu và thay đổi cơ cấu hoặc so sánh cũng như biểu hiện mối liên hệ giữa các hiện tượng.

Ví dụ: Biểu diễn số lượng cán bộ khoa học công nghệ chia theo nam nữ của 4 năm: 2000, 2001, 2002 và 2003 qua biểu đồ 3.2.1.

Biểu đồ 3.2.1: Hình cột phản ánh số lượng cán bộ khoa học công nghệ



Đồ thị trên vừa phản ánh quá trình phát triển của cán bộ KHCN vừa so sánh cũng như phản ánh mối liên hệ giữa cán bộ là nam và nữ.

3.2.2. Biểu đồ diện tích

Biểu đồ diện tích là loại biểu đồ, trong đó các thông tin thống kê được biểu hiện bằng các loại diện tích hình học như hình vuông, hình chữ nhật, hình tròn, hình ô van,...

Biểu đồ diện tích thường được dùng để biểu hiện kết cấu và biến

động cơ cấu của hiện tượng.

Tổng diện tích của cả hình là 100%, thì diện tích từng phần tương ứng với mỗi bộ phận phản ánh cơ cấu của bộ phận đó.

Biểu đồ diện tích hình tròn còn có thể biểu hiện được cả cơ cấu, biến động cơ cấu kết hợp thay đổi mức độ của hiện tượng. Trong trường hợp này số đo của góc các hình quạt phản ánh cơ cấu và biến động cơ cấu, còn diện tích toàn hình tròn phản ánh quy mô của hiện tượng.

Khi vẽ đồ thị ta tiến hành như sau:

- Lấy giá trị của từng bộ phận chia cho giá trị chung của chỉ tiêu nghiên cứu để xác định tỷ trọng (%) của từng bộ phận đó. Tiếp tục lấy 360 c (360⁰) chia cho 100 rồi nhân với tỷ trọng của từng bộ phận sẽ xác định được góc độ tương ứng với cơ cấu của từng bộ phận.

- Xác định bán kính của mỗi hình tròn có diện tích tương ứng là $S: R = \sqrt{S : \pi}$ vì diện tích hình tròn: $S = \pi.R^2$. Khi có độ dài của bán kính mỗi hình tròn, ta sẽ dễ dàng vẽ được các hình tròn đó.

Ví dụ: Có số lượng về học sinh phổ thông phân theo cấp học 3 năm 2001, 2002 và 2003 như bảng 3.2.1:

Bảng 3.2.1: Học sinh phổ thông phân theo cấp học

	2001		2002		2003	
	Số lượng (Người)	Cơ cấu (%)	Số lượng (Người)	Cơ cấu (%)	Số lượng (Người)	Cơ cấu (%)
Tổng số học sinh	1000	100,0	1140	100,0	1310	100,0
<i>Chia ra:</i>						
Tiểu học	500	50,0	600	53,0	700	53,5
Trung học cơ sở	300	30,0	320	28,0	360	27,5
Trung học phổ thông	200	20,0	220	19,0	250	19,0

Từ số liệu bảng 3.2.1 ta tính các bán kính tương ứng:

$$\text{Năm 2001: } R = \sqrt{1000 / 3,14} = 17,84$$

$$\text{Năm 2002: } R = \sqrt{1140 / 3,14} = 19,05$$

$$\text{Năm 2003: } R = \sqrt{1310 / 3,14} = 20,42$$

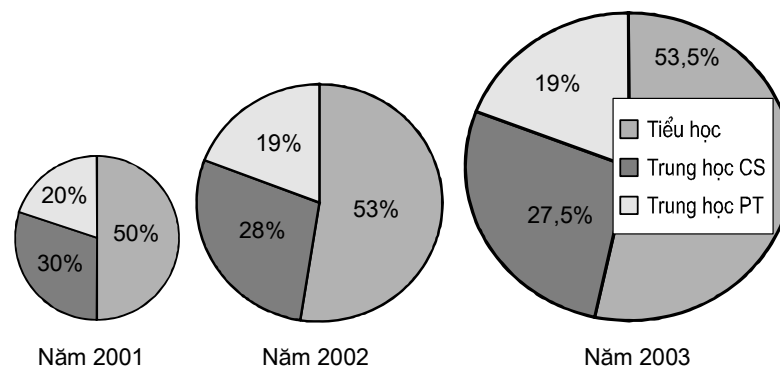
Nếu năm 2001 lấy $R = 1,00$

Thì năm 2002 có $R = 19,05 : 17,84 = 1,067$

Năm 2003 có $R = 20,42 : 17,84 = 1,144$

Ta vẽ các hình tròn tương ứng với 3 năm (2001, 2002 và 2003) có bán kính là 1,00; 1,067 và 1,144 rồi mỗi hình tròn chia diện tích các hình tròn theo cơ cấu học sinh các cấp tương ứng như số liệu ở bảng trên lên các biểu đồ. Kết quả 3 hình tròn được vẽ phản ánh cả quy mô học sinh phổ thông lẫn cơ cấu và biến động cơ cấu theo cấp học của học sinh qua các năm 2001, 2002 và 2003.

Biểu đồ 3.2.2: Biểu đồ diện tích hình tròn phản ánh số lượng và cơ cấu học sinh phổ thông



3.2.3. Biểu đồ tượng hình

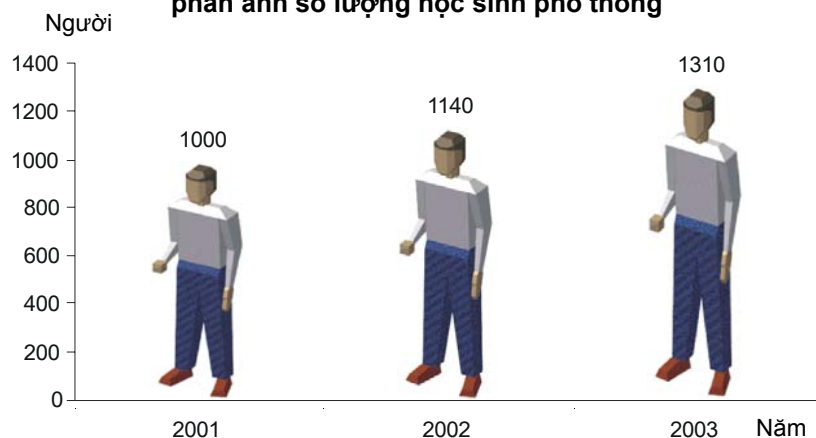
Biểu đồ tượng hình là loại đồ thị thống kê, trong đó các tài liệu

thống kê được thể hiện bằng các hình vẽ tượng trưng. Biểu đồ tượng hình được dùng rộng rãi trong việc tuyên truyền, phổ biến thông tin trên các phương tiện sử dụng rộng rãi. Biểu đồ hình tượng có nhiều cách vẽ khác nhau, tùy theo sáng kiến của người trình bày mà lựa chọn loại hình vẽ tượng hình cho phù hợp và hấp dẫn.

Tuy nhiên khi sử dụng loại biểu đồ này phải theo nguyên tắc: cùng một chỉ tiêu phải được biểu hiện bằng cùng một loại hình vẽ, còn chỉ tiêu đó ở các trường hợp nào có trị số lớn nhỏ khác nhau thì sẽ biểu hiện bằng hình vẽ có kích thước lớn nhỏ khác nhau theo tỷ lệ tương ứng.

Trở lại ví dụ trên số lượng học sinh phổ thông được biểu diễn bằng các cậu bé cấp sáu, năm 2002 có số lượng lớn hơn năm 2001 và năm 2003 có số lượng lớn hơn năm 2002 thì cậu bé ứng với năm 2002 phải lớn hơn cậu bé ứng với năm 2001 và cậu bé ứng với năm 2003 phải lớn hơn cậu bé ứng với năm 2002 (xem biểu đồ 3.2.3).

Biểu đồ 3.2.3: Biểu đồ tượng hình, phản ánh số lượng học sinh phổ thông



3.2.4. Đồ thị đường gấp khúc

Đồ thị đường gấp khúc là loại đồ thị thống kê biểu hiện các tài

liệu bằng một đường gấp khúc nối liền các điểm trên một hệ tọa độ, thường là hệ tọa độ vuông góc.

Đồ thị đường gấp khúc được dùng để biểu hiện quá trình phát triển của hiện tượng, biểu hiện tình hình phân phối các đơn vị tổng thể theo một tiêu thức nào đó, hoặc biểu thị tình hình thực hiện kế hoạch theo từng thời gian của các chỉ tiêu nghiên cứu.

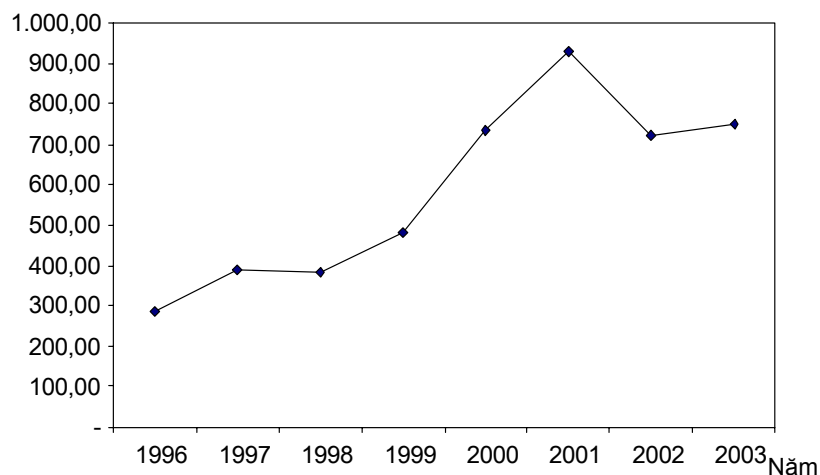
Trong một đồ thị đường gấp khúc, trục hoành thường được biểu thị thời gian, trục tung biểu thị mức độ của chỉ tiêu nghiên cứu. Cũng có khi các trục này biểu thị hai chỉ tiêu có liên hệ với nhau, hoặc lượng biến và các tần số (hay tần suất) tương ứng. Độ phân chia trên các trục cần được xác định cho thích hợp vì có ảnh hưởng trực tiếp đến độ dốc của đồ thị. Mặt khác, cần chú ý là trên mỗi trục tọa độ chiều dài của các khoảng phân chia tương ứng với sự thay đổi về lượng của chỉ tiêu nghiên cứu phải bằng nhau.

Ví dụ: Sản lượng cà phê xuất khẩu của Việt Nam qua các năm từ 1996 đến 2003 (nghìn tấn) có kết quả như sau: 283,3; 391,6; 382,0; 482,0; 733,9; 931,0; 722, 0 và 749,0.

Số liệu trên được biểu diễn qua đồ thị đường gấp khúc 3.2.4.

Đồ thị 3.2.4: Đường gấp khúc phản ánh biến động của sản lượng cà phê xuất khẩu qua các năm của Việt Nam

Nghìn tấn



3.2.5. Biểu đồ hình màng nhện

Biểu đồ hình màng nhện là loại đồ thị thống kê dùng để phản ánh kết quả đạt được của hiện tượng lặp đi lặp lại về mặt thời gian, ví dụ phản ánh về biến động thời vụ của một chỉ tiêu nào đó qua 12 tháng trong năm. Để lập đồ thị hình màng nhện ta vẽ một hình tròn bán kính R , sao cho R lớn hơn trị số lớn nhất của chỉ tiêu nghiên cứu (lớn hơn bao nhiêu lần không quan trọng, miễn là đảm bảo tỷ lệ nào đó để hình vẽ được cân đối, kết quả biểu diễn của đồ thị dễ nhận biết). Sau đó chia đường tròn bán kính R thành các phần đều nhau theo số kỳ nghiên cứu (ở đây là 12 tháng) bởi các đường thẳng đi qua tâm đường tròn. Nối các giao điểm của bán kính cắt đường tròn ta được đa giác đều nội tiếp đường tròn. Đó là giới hạn phạm vi của đồ thị. Độ dài đo từ tâm đường tròn đến các điểm xác định theo các đường phân chia đường tròn nói trên chính là các đại lượng cần biểu hiện của hiện tượng tương ứng với mỗi thời kỳ. Nối các điểm xác định sẽ được hình vẽ của đồ thị hình màng nhện.

Ví dụ: Có số liệu về trị giá xuất, nhập khẩu hải sản của tỉnh "X" 2 năm (2002 và 2003) như sau:

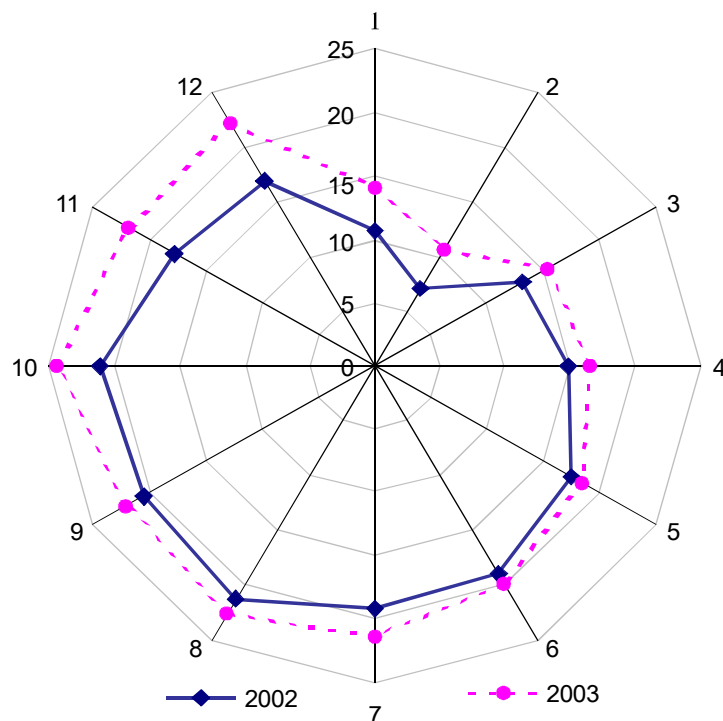
Bảng 3.2.2: Giá trị xuất khẩu hải sản trong 12 tháng của năm 2002 và 2003

ĐVT: Triệu đồng

NămN Tháng	2002	2003	NămN Tháng	2002	2003	NămN Tháng	2002	2003
A	1	2	A	1	2	A	1	2
1	10,7	14,0	5	17,4	18,4	9	20,5	22,2
2	7,0	10,5	6	18,9	19,8	10	21,1	24,4
3	13,1	15,4	7	19,1	21,3	11	17,7	21,8
4	14,8	16,5	8	21,2	22,5	12	16,8	22,1

Từ số liệu ta nhận thấy tháng 10 năm 2003 tỉnh "X" có trị giá xuất khẩu lớn nhất (24, 4 triệu USD). Ta xem 1 triệu USD là một đơn vị và sẽ vẽ đường tròn có bán kính $R = 25 > 24,4$ đơn vị. Chia đường tròn thành 12 phần đều nhau, vẽ các đường thẳng tương ứng cắt đường tròn tại 12 điểm. Nối các điểm lại có đa giác đều 12 cạnh nội tiếp đường tròn. Căn cứ số liệu của bảng ta xác định các điểm tương ứng với giá trị xuất khẩu đạt được của các tháng trong từng năm rồi nối các điểm đó lại thành đường liền ta được đồ thị hình màng nhện biểu diễn kết quả xuất khẩu qua các tháng trong 2 năm của tỉnh "X" (xem đồ thị 3.2.5).

Đồ thị 3.2.5. Đồ thị hình màng nhện về kết quả xuất khẩu



Sự mô tả của đồ thị hình màng nhện cho phép ta quan sát và so sánh không chỉ kết quả xuất khẩu giữa các tháng khác nhau trong cùng một năm, mà cả kết quả sản xuất giữa các tháng cùng tên của các năm khác nhau cũng như xu thế biến động chung về xuất khẩu của các năm.

3.3. PHƯƠNG PHÁP PHÂN TÍCH DẪY SỐ BIẾN ĐỘNG THEO THỜI GIAN

3.3.1. Khái niệm và đặc điểm của dãy số biến động theo thời gian

Dãy số biến động theo thời gian (còn gọi là dãy số động thái) là dãy các trị số của một chỉ tiêu thống kê được sắp xếp theo thứ tự thời

gian, dùng để phản ánh quá trình phát triển của hiện tượng. Ví dụ sản lượng điện Việt Nam (tỷ kw /h) từ 1995 đến 2002 như sau: 14,7; 17,0; 19,3; 21,7; 23,6; 26,6; 30,7; 35,6.

Trong dãy số biến động theo thời gian có hai yếu tố: thời gian và chỉ tiêu phản ánh hiện tượng nghiên cứu. Thời gian trong dãy số có thể là ngày, tháng, năm,... tùy mục đích nghiên cứu; chỉ tiêu phản ánh hiện tượng nghiên cứu có thể biểu hiện bằng số tuyệt đối, số tương đối hay số bình quân.

Căn cứ vào tính chất của thời gian trong dãy số có thể phân biệt hai loại:

+ Dãy số biến động theo thời kỳ (gọi tắt là dãy số thời kỳ): Dãy số trong đó các mức độ của chỉ tiêu biểu hiện mặt lượng của hiện tượng trong một khoảng thời gian nhất định. Ví dụ: Dãy số về sản lượng điện sản xuất ra hàng năm; GDP tính theo giá so sánh thời kỳ 1990 - 2002,...

+ Dãy số biến động theo thời điểm (gọi tắt là dãy số thời điểm): Dãy số trong đó các mức độ của chỉ tiêu biểu hiện mặt lượng của hiện tượng ở những thời điểm nhất định. Ví dụ: Dãy số về số học sinh phổ thông nhập học có đến ngày khai giảng hàng năm,...

Căn cứ vào đặc điểm của dãy số biến động theo thời gian ta có thể vạch rõ xu hướng, tính quy luật phát triển của hiện tượng theo thời gian và từ đó có thể dự đoán khả năng hiện tượng có thể xảy ra trong tương lai.

Các trị số của chỉ tiêu trong dãy số thời gian phải thống nhất về nội dung; phương pháp và đơn vị tính; thống nhất về khoảng cách thời gian và phạm vi không gian nghiên cứu của hiện tượng để bảo đảm tính so sánh được với nhau.

3.3.2. Các chỉ tiêu phân tích dãy số biến động theo thời gian

3.3.2.1. Mức độ bình quân theo thời gian

Mức độ bình quân theo thời gian là số bình quân về các mức độ của chỉ tiêu trong dãy số thời gian, biểu hiện mức độ điển hình của hiện tượng nghiên cứu trong một khoảng thời gian dài với công thức tính như sau:

a. Mức độ bình quân theo thời gian tính từ một dãy số thời kỳ.

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i \quad ; \quad (3.3.1a)$$

Trong đó:

$\bar{y}$ - Mức độ bình quân theo thời gian;

y_i ($i = 1, 2, 3, \dots, n$) - Các mức độ của chỉ tiêu trong dãy số thời kỳ;

n - Số thời kỳ trong dãy số.

b. Mức độ bình quân theo thời gian tính từ một dãy số thời điểm

$$\bar{y} = \frac{\frac{y_1}{2} + y_2 + \dots + y_{n-1} + \frac{y_n}{2}}{n-1} \quad ; \quad (3.3.1b)$$

Trong đó:

$y_1, y_2, \dots, y_n$ - Các mức độ của chỉ tiêu trong dãy số thời điểm;

n - Số thời điểm trong dãy số.

- Nếu dãy số thời điểm có khoảng cách thời gian không đều nhau, phải lấy thời gian trong mỗi khoảng cách làm quyền số.

$$\bar{y} = \frac{\sum y_i t_i}{\sum t_i}$$

Trong đó: t_i - Thời gian trong mỗi khoảng cách.

3.3.2.2. Lượng tăng tuyệt đối

Lượng tăng tuyệt đối là hiệu số giữa hai mức độ của chỉ tiêu trong dãy số thời gian, phản ánh sự thay đổi của mức độ hiện tượng qua hai

thời gian khác nhau. Nếu hướng phát triển của hiện tượng tăng thì lượng tăng tuyệt đối mang dấu dương và ngược lại. Tùy theo mục đích nghiên cứu có thể tính các lượng tăng tuyệt đối sau:

a. *Lượng tăng tuyệt đối liên hoàn* (hay lượng tăng tuyệt đối từng kỳ). Đó là hiệu số của một mức độ nào đó trong dãy số ở kỳ nghiên cứu với mức độ của kỳ kề liền trước nó. Công thức tính như sau:

$$\delta_i = y_i - y_{i-1} \quad ; \quad (3.3.2a)$$

Trong đó: δ_i - Lượng tăng tuyệt đối liên hoàn;

y_i - Mức độ của chỉ tiêu trong dãy số kỳ nghiên cứu;

y_{i-1} - Mức độ ở kỳ kề liền trước mức độ kỳ nghiên cứu.

b. *Lượng tăng tuyệt đối định gốc* (hay lượng tăng tuyệt đối cộng dồn). Đó là hiệu số giữa mức độ nào đó ở kỳ nghiên cứu trong dãy số với mức độ được chọn làm gốc không thay đổi (thường là mức độ đầu tiên trong dãy số). Công thức tính:

$$\Delta_i = y_i - y_1 \quad ; \quad (3.3.2b)$$

Trong đó:

Δ_i - Lượng tăng tuyệt đối định gốc;

y_i - Mức độ của chỉ tiêu trong dãy số kỳ nghiên cứu;

y_1 - Mức độ của chỉ tiêu ở kỳ được chọn làm gốc so sánh.

c. *Lượng tăng tuyệt đối bình quân*. Đó là số bình quân của các lượng tăng tuyệt đối từng kỳ. Công thức tính:

$$\bar{\delta} = \frac{\sum_{i=2}^n \delta_i}{n-1} = \frac{\Delta_n}{n-1} = \frac{y_n - y_1}{n-1} \quad ; \quad (3.3.2c)$$

Trong đó: $\bar{\delta}$ - Lượng tăng tuyệt đối bình quân.

3.3.2.3. Tốc độ phát triển (Chỉ số phát triển)

Tốc độ phát triển là chỉ tiêu tương đối dùng để phản ánh nhịp

điều biến động của hiện tượng nghiên cứu qua hai thời kỳ / thời điểm khác nhau và được biểu hiện bằng số lần hay số phần trăm. Tốc độ phát triển được tính bằng cách so sánh giữa hai mức độ của chỉ tiêu trong dãy số biến động theo thời gian, trong đó một mức độ được chọn làm gốc so sánh. Tùy theo mục đích nghiên cứu, có thể tính các loại tốc độ phát triển sau:

a. Tốc độ phát triển liên hoàn (hay tốc độ phát triển từng kỳ): Dùng để phản ánh sự phát triển của hiện tượng qua từng thời gian ngắn liền nhau, được tính bằng cách so sánh một mức độ nào đó trong dãy số ở kỳ nghiên cứu với mức độ liền trước đó. Công thức tính:

$$t_i = \frac{y_i}{y_{i-1}} \quad ; \quad (3.3.3a)$$

Trong đó:

t_i - Tốc độ phát triển liên hoàn;

y_i - Mức độ của chỉ tiêu trong dãy số ở kỳ nghiên cứu;

y_{i-1} - Mức độ của chỉ tiêu ở kỳ liền kề trước kỳ nghiên cứu.

b. Tốc độ phát triển định gốc (hay tốc độ phát triển cộng dồn): Dùng để phản ánh sự phát triển của hiện tượng qua một thời gian dài, được tính bằng cách so sánh mức độ nào đó của kỳ nghiên cứu trong dãy số với mức độ được chọn làm gốc không thay đổi (thường là mức độ đầu tiên trong dãy số). Công thức tính:

$$T_i = \frac{y_i}{y_1} \quad ; \quad (3.3.3b)$$

Trong đó:

T_i - Tốc độ phát triển định gốc;

y_i - Mức độ của chỉ tiêu của kỳ nghiên cứu;

y_1 - Mức độ của chỉ tiêu được chọn làm gốc so sánh.

Tốc độ phát triển định gốc bằng tích số các tốc độ phát triển liên

hoàn, mỗi liên hệ này được viết dưới dạng công thức như sau:

$$T_i = t_2 \times t_3 \times \dots \times t_n = \prod_{i=2}^n t_i$$

c. Tốc độ phát triển bình quân: Dùng để phản ánh nhịp độ phát triển điển hình của hiện tượng nghiên cứu trong một thời gian dài, được tính bằng số bình quân nhân của các tốc độ phát triển liên hoàn. Chỉ tiêu tốc độ phát triển bình quân chỉ có ý nghĩa đối với những hiện tượng phát triển tương đối đều đặn theo một chiều hướng nhất định. Công thức tính như sau:

$$\bar{t} = \sqrt[n]{t_2 \times t_3 \times \dots \times t_n} = \sqrt[n]{\prod_{i=2}^n t_i} = \sqrt[n]{T_n} \quad ; \quad (3.3.3c)$$

Trong đó:

$\bar{t}$ - Tốc độ phát triển bình quân;

t_i ($i = 2, 3, \dots, n$) - Các tốc độ phát triển liên hoàn tính được từ một dãy số biến động theo thời gian gồm n mức độ.

Ví dụ: Từ số liệu về sản lượng điện của Việt Nam thời kỳ 1995 - 2002, ký hiệu i bằng 1 đối với năm 1995 và i bằng 8 đối với năm 2002, tính được tốc độ phát triển bình quân như sau:

- Tốc độ phát triển định gốc (2002 so với 1995):

$$T_{8/1} = \frac{35,6}{17,7} = 2,482 \text{ hoặc } 248,2\%$$

- Tốc độ phát triển bình quân thời kỳ 1995 - 2002:

$$\bar{t} = \sqrt[8]{2,482} = 1,139 \text{ hoặc } 113,9\%$$

3.3.2.4. Tốc độ tăng

Tốc độ tăng là chỉ tiêu tương đối phản ánh nhịp điệu tăng /giảm của hiện tượng qua thời gian và biểu hiện bằng số lần hoặc số phần trăm, được tính bằng cách so sánh lượng tăng tuyệt đối giữa hai thời

kỳ với mức độ kỳ gốc chọn làm căn cứ so sánh. Tùy theo mục đích nghiên cứu có thể tính các loại tốc độ tăng sau:

a. Tốc độ tăng liên hoàn (từng kỳ)

$$i_i = \frac{y_i - y_{i-1}}{y_{i-1}} = \frac{\delta_i}{y_{i-1}} \quad ; \quad (3.3.4a)$$

Trong đó:

i_i - Tốc độ tăng liên hoàn;

δ_i - Lượng tăng tuyệt đối liên hoàn;

y_i - Mức độ của chỉ tiêu kỳ nghiên cứu;

y_{i-1} - Mức độ của chỉ tiêu trước kỳ nghiên cứu.

b. Tốc độ tăng định gốc (cộng dồn)

$$I_i = \frac{y_i - y_1}{y_1} = \frac{\Delta_i}{y_1} \quad ; \quad (3.3.4b)$$

Trong đó:

I_i - Tốc độ tăng định gốc;

Δ_i - Lượng tăng tuyệt đối định gốc.

Mối liên hệ giữa tốc độ phát triển và tốc độ tăng như sau:

Nếu tính bằng số lần: Tốc độ tăng = Tốc độ phát triển – 1

Nếu tính bằng phần trăm: Tốc độ tăng = Tốc độ phát triển – 100.

c. Tốc độ tăng bình quân phản ánh nhịp độ tăng điển hình của hiện tượng nghiên cứu trong thời gian dài.

Tốc độ tăng bình quân ($\bar{i}$) = Tốc độ phát triển bình quân ($\bar{t}$) – 1 (hay 100).

Từ kết quả tính tốc độ phát triển bình quân năm về điện sản xuất ra: $\bar{t} = 1,139$ hoặc $113,9\%$, tính được tốc độ tăng bình quân ($\bar{i}$) thời kỳ 1995-2002:

$$\bar{i} = 1,139 - 1 = 0,139$$

hoặc $\bar{i} = 113,9 - 100 = 13,9\%$

3.3.2.5. Giá trị tuyệt đối của 1% tăng lên

Giá trị tuyệt đối của 1% tăng lên nói lên mức độ thực tế của 1% tốc độ tăng, được tính bằng cách đem chia lượng tuyệt đối từng kỳ cho tốc độ tăng từng kỳ. Công thức tính:

$$\text{Giá trị tuyệt đối của 1\% tăng lên} = \frac{\text{Lượng tăng tuyệt đối từng kỳ}}{\text{Tốc độ tăng từng kỳ (\%)}} \quad ; \quad (3.3.5a)$$

hoặc:

$$\text{Giá trị tuyệt đối của 1\% tăng lên} = \frac{\text{Mức độ kỳ gốc (liên hoàn)}}{100} \quad ; \quad (3.3.5b)$$

Ví dụ: Sản lượng điện của Việt Nam năm 2001 ($i=7$) là 30,7 tỷ kwh, năm 2002 ($i=8$) là 35,6 tỷ kwh. Như vậy, tính được các chỉ tiêu năm 2002 so với năm 2001.

- Lượng tăng tuyệt đối:

$$\delta_{8/7} = 35,6 - 30,7 = 4,9 \text{ (tỷ kwh)}$$

- Tốc độ tăng:

$$i_{8/7} = \frac{4,9}{30,7} = 0,1596 \text{ hoặc } 15,96\%$$

- Giá trị tuyệt đối của 1% sản lượng điện tăng lên:

$$a_{8/7} = \frac{4,9}{15,96} = \frac{30,7}{100} = 0,307 \text{ (tỷ kwh)}$$

3.3.3. Một số phương pháp biểu hiện xu hướng biến động cơ bản của hiện tượng

3.3.3.1. Phương pháp mở rộng khoảng cách thời gian

Đó là phương pháp điều chỉnh một dãy số biến động theo thời gian, nhằm nêu lên xu hướng phát triển cơ bản của hiện tượng. Phương pháp này được áp dụng khi dãy số có những khoảng thời gian ngắn và có quá nhiều mức độ, do đó không thể hiện được rõ xu hướng phát triển của hiện tượng. Có thể rút bớt các mức độ trong dãy số bằng cách mở rộng các khoảng cách thời gian của các mức độ, như biến đổi mức độ chỉ tiêu hàng ngày thành mức độ chỉ tiêu hàng tháng, từ hàng tháng thành quý, từ hàng quý thành hàng năm,...

3.3.3.2. Phương pháp số bình quân trượt

Đó là phương pháp điều chỉnh một dãy số biến động theo thời gian có các mức độ lên xuống thất thường, nhằm loại trừ các nhân tố ngẫu nhiên và phát hiện xu hướng phát triển cơ bản của hiện tượng. Áp dụng phương pháp này, trước hết người ta lấy một nhóm (ba, bốn, năm,...) mức độ đầu tiên để tính một số bình quân. Tiếp tục tính các số bình quân trượt của các nhóm khác bằng cách lần lượt bỏ mức độ trên cùng và thêm vào mức độ kế tiếp cho đến mức độ cuối cùng của nhóm.

Ví dụ: Một dãy số biến động theo thời gian gồm các mức độ $y_1, y_2, \dots, y_n$. Tính số bình quân di động cho từng nhóm 3 mức độ.

$$\begin{aligned}\bar{y}_I &= \frac{y_1 + y_2 + y_3}{3}; & \bar{y}_{II} &= \frac{y_2 + y_3 + y_4}{3}; \\ \bar{y}_{III} &= \frac{y_3 + y_4 + y_5}{3}; & \dots & \end{aligned} \quad (3.3.6)$$

Như vậy cuối cùng cũng có thể lập một dãy số mới gồm các số bình quân di động $\bar{y}_I, \bar{y}_{II}, \bar{y}_{III}, \dots$ có thể tiếp tục điều chỉnh một vài lần nữa, bằng cách tính số bình quân di động của các số bình quân di động trong dãy số.

3.3.3.3. Phương pháp điều chỉnh bằng phương trình toán học

Phương pháp điều chỉnh bằng phương trình toán học các mức độ của chỉ tiêu trong *một dãy số biến động theo thời gian*, nhằm nêu lên xu hướng phát triển cơ bản hiện tượng. Theo phương pháp này, có thể căn cứ vào tính chất biến động của các mức độ của chỉ tiêu trong dãy số để xác định một phương trình hồi quy biểu diễn biến động theo đường thẳng hoặc đường cong, từ đó tính các mức độ lý thuyết thay cho các mức độ thực tế của chỉ tiêu. Bằng phương pháp bình phương nhỏ nhất xây dựng được hệ thống phương trình chuẩn tắc để tính các tham số của các phương trình cần điều chỉnh.

Sau đây là một số dạng phương trình hồi quy đơn giản thường được sử dụng:

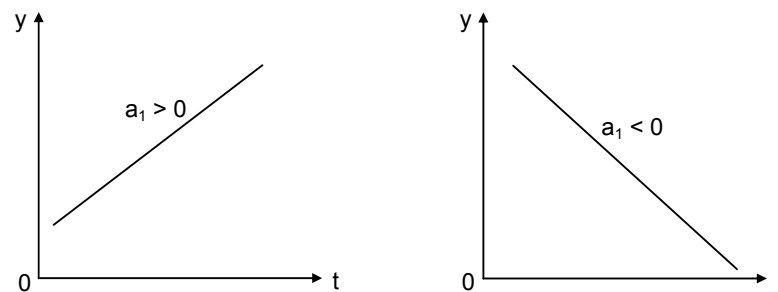
* *Phương trình đường thẳng*

$$\bar{y}_t = a_0 + a_1 t \quad ; \quad (3.3.7a)$$

Các tham số a_0 và a_1 được xác định theo hệ phương trình chuẩn tắc sau đây:

$$\begin{cases} \sum y = na_0 + a_1 \sum t \\ \sum yt = a_0 \sum t + a_1 \sum t^2 \end{cases} \quad ; \quad (3.3.7b)$$

Đồ thị biểu diễn phương trình đường thẳng ($y = a_0 + a_1 t$) có dạng:



* *Phương trình parabol bậc 2*

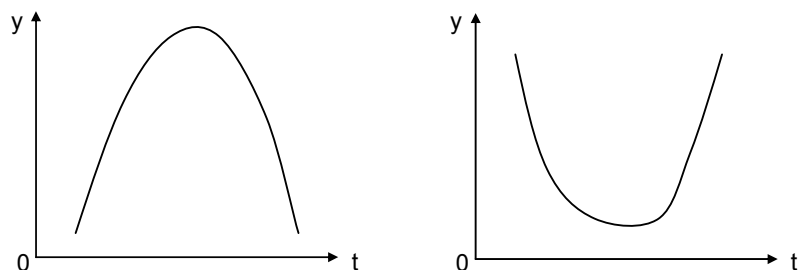
$$\bar{y}_t = a_0 + a_1 t + a_2 t^2 \quad ; \quad (3.3.8a)$$

Các tham số a_0, a_1 và a_2 được xác định theo hệ phương trình

chuẩn tắc sau đây:

$$\begin{cases} \Sigma y = na_0 + a_1 \Sigma t + a_2 \Sigma t^2 \\ \Sigma yt = a_0 \Sigma t + a_1 \Sigma t^2 + a_2 \Sigma t^3 \\ \Sigma yt^2 = a_0 \Sigma t^2 + a_1 \Sigma t^3 + a_2 \Sigma t^4 \end{cases} ; \quad (3.3.8b)$$

Đồ thị biểu diễn phương trình đường bậc hai ($y = a_0 + a_1 t + a_2 t^2$) có dạng:



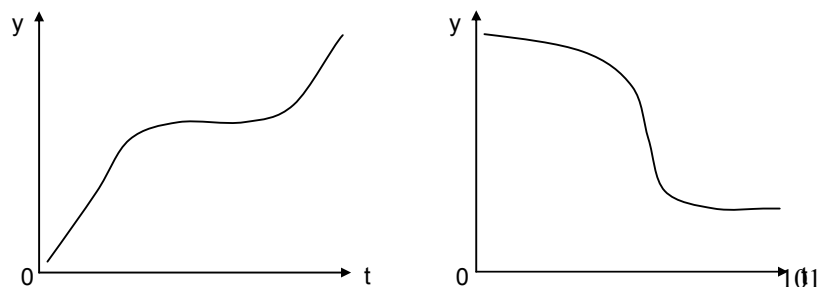
* Phương trình bậc 3

$$\bar{y}_t = a_0 + a_1 t + a_2 t^2 + a_3 t^3 ; \quad (3.3.9a)$$

Các tham số a_0 , a_1 , a_2 và a_3 của phương trình bậc ba được xác định theo hệ phương trình chuẩn tắc sau:

$$\begin{cases} \Sigma y = na_0 + a_1 \Sigma t + a_2 \Sigma t^2 + a_3 \Sigma t^3 \\ \Sigma yt = a_0 \Sigma t + a_1 \Sigma t^2 + a_2 \Sigma t^3 + a_3 \Sigma t^4 \\ \Sigma yt^2 = a_0 \Sigma t^2 + a_1 \Sigma t^3 + a_2 \Sigma t^4 + a_3 \Sigma t^5 \\ \Sigma yt^3 = a_0 \Sigma t^3 + a_1 \Sigma t^4 + a_2 \Sigma t^5 + a_3 \Sigma t^6 \end{cases} ; \quad (3.3.9b)$$

Đồ thị biểu diễn phương trình đường bậc ba có dạng:



* Phương trình hàm mũ

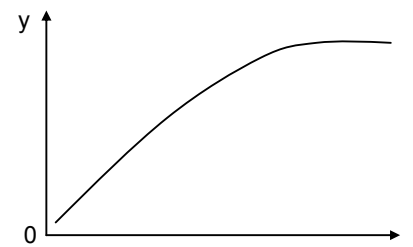
$$\bar{y}_t = a_0 a_1^t ; \quad (3.3.10a)$$

Phương trình hàm số mũ được sử dụng khi các tốc độ phát triển liên hoàn xấp xỉ nhau.

Các tham số a_0 và a_1 được xác định theo hệ phương trình chuẩn tắc sau đây:

$$\begin{cases} \Sigma \lg y = n \lg a_0 + \lg a_1 \Sigma t \\ \Sigma t. \lg y = \lg a_0 \Sigma t + \lg a_1 \Sigma t^2 \end{cases} ; \quad (3.3.10b)$$

Đồ thị biểu diễn phương trình hàm số mũ có dạng:



Xét một ví dụ đơn giản sau đây điều chỉnh theo phương trình đường thẳng ($\bar{y}_t = a_0 + a_1 t$): Giả sử có tài liệu về năng suất lúa bình quân một vụ của một địa phương qua một số năm và lập thành bảng tính 3.3.1 như sau:

Bảng 3.3.1: Bảng tính toán các tham số của hệ phương trình chuẩn tắc

Năm	Năng suất bình quân (Tạ/ha) (y)	Phần tính toán			
		Thứ tự thời gian (t)	.t ²	.ty	$\bar{y}_t$

1998	30	1	1	30	30,4
1999	32	2	4	64	31,2
2000	31	3	9	93	32,0
2001	34	4	16	136	32,8
2002	33	5	25	165	33,6
Cộng	160	15	55	488	

Dựa vào hệ phương trình 3.3.7b nêu trên, thay các số liệu tính toán được trong bảng vào hệ phương trình, có:

$$\begin{cases} 160 = 5a_0 + a_1 15 \\ 488 = a_0 15 + a_1 55 \end{cases}$$

Giải ra ta được: $\begin{cases} a_0 = 29,6 \\ a_1 = 0,8 \end{cases}$

Từ đó: $\bar{y}_t = 29,6 + 0,8t$

$a_1 = 0,8$ phản ánh mức tăng bình quân hàng năm của năng suất lúa là 0,8 t/ha.

Để điều chỉnh dãy số biến động theo hàm số phù hợp với thực tế, trước hết phải dựa vào lý thuyết kinh tế để phân tích tính chất và xu thế biến động của hiện tượng. Sau đó dựa vào số liệu thực tế đưa lên đồ thị để nhận biết dạng hàm, từ đó chọn một số dạng cơ bản phù hợp để điều chỉnh, thay giá trị thời gian t vào các hàm đã điều chỉnh để tính các giá trị lý thuyết của từng hàm ($\hat{y}_t$). Mỗi phương trình điều chỉnh sẽ tính được một hệ số mô tả.

$$V_{\bar{y}} = \frac{\sigma_y}{\bar{y}} \times 100 \quad ; \quad (3.3.11)$$

Trong đó:

$$\sigma_y = \sqrt{\frac{\sum_{t=1}^n (y_t - \bar{y}_t)^2}{n}} \quad \text{và} \quad \bar{y} = \frac{\sum_{t=1}^n y_t}{n}$$

Phương trình nào có hệ số mô tả nhỏ nhất, tức là hệ số xác định lớn nhất thì sẽ phản ánh phù hợp nhất xu thế biến động của chỉ tiêu và đó là phương trình điều chỉnh được lựa chọn.

Việc giải các hệ phương trình chuẩn tắc nêu trên để tính các tham số $a_0, a_1, a_2, \dots$ cũng như tính toán các giá trị lý thuyết ($\hat{y}_t$) theo các mô hình hồi quy khá phức tạp và có khối lượng tính toán khá lớn. Nhưng ngày nay nhờ công cụ máy tính, chúng ta có thể thực hiện được các yêu cầu đó một cách nhanh chóng và thuận lợi. Các kết quả của bài toán máy tính chạy ra còn cho ta những kết quả về hệ số xác định, hệ số mô tả để có căn cứ kết luận mức độ đại diện của từng đường hồi quy lý thuyết làm cơ sở cho ta lựa chọn mô hình tốt nhất.

3.3.3.4. Phân tích biến động thời vụ

Đó là phương pháp nghiên cứu và xác định sự biến động một cách có quy luật vào những thời kỳ nhất định trong vòng một năm của một hiện tượng kinh tế - xã hội. Biến động thời vụ có thể do những nguyên nhân như điều kiện địa lý, thời tiết, tập quán sinh hoạt của con người, ... Ví dụ: Trong công nghiệp, tình hình chế biến chè, mía, hoa quả hộp, ... phụ thuộc vào vụ thu hoạch; trong xây dựng cơ bản khối lượng xây lắp bị ảnh hưởng bởi thời tiết trong năm; trong thương nghiệp nhiều mặt hàng có lượng tiêu thụ nhiều hay ít tùy theo mùa.

Biến động thời vụ ảnh hưởng nhiều đến tình hình sản xuất và sinh hoạt, nhiệm vụ của thống kê khi phân tích biến động thời vụ là: Dựa trên số liệu thống kê nhiều năm (ít nhất là 3 năm) tính các chỉ số thời vụ.

* Trường hợp biến động thời vụ của các tháng tương ứng giữa các năm tương đối ổn định, không có hiện tượng tăng (hoặc giảm) rõ rệt thì chỉ số thời vụ được tính theo công thức sau đây:

$$I_i = \frac{\bar{y}_i}{\bar{y}_0} \quad ; \quad (3.3.12a)$$

Trong đó:

I_i - Chỉ số thời vụ của thời gian t ;

$\bar{y}_i$ - Số bình quân các mức độ của các thời gian cùng tên i ;

$\bar{y}_0$ - Số bình quân của tất cả các mức độ trong các dãy số.

Ví dụ: Có tài liệu về mức tiêu thụ hàng hoá "X" ở một địa phương trong 3 năm như bảng 3.3.2:

Bảng 3.3.2: Tính toán chỉ số thời vụ

Năm N (j) Tháng (i)	Mức tiêu thụ hàng hoá "X" (y_{it} - triệu đồng)			$\bar{y}_i$	$\frac{\bar{y}_i}{\bar{y}_0} 100$
	2000	2001	2002		
1	2	3	4	5	6
1	1495	1500	1490	1495	62,9
2	1461	1490	1480	1477	62,2
3	1533	1599	1604	1578	66,4
4	1922	2210	2005	2046	86,1
5	2746	2804	2745	2765	116,4
6	3289	3282	3250	3274	137,8
7	3523	3620	3700	3614	152,1
8	3330	3300	3215	3282	138,2
9	2597	2604	2599	2597	109,3
10	2249	2205	2304	2253	94,8
11	2144	2200	2190	2178	91,7
12	1983	1889	1950	1941	81,7
Tổng cả năm	28272	28703	28523	2375	

Từ số liệu cột 2, 3, 4 bảng 3.3.2, ta tính mức tiêu thụ hàng hoá bình quân tháng 1: $\bar{y}_1 = \frac{1495 + 1500 + 1490}{3} = 1495$ triệu đồng.

Bằng cách tương tự ta tính giá trị trung bình tháng 2, 3,..., 12 như cột 5 của bảng.

Tiếp tục tính bình quân chung của tất cả các mức độ là:

$$\bar{y}_0 = \frac{\sum_{j=1}^3 \sum_{i=1}^{12} y_{ij}}{36} = \frac{\sum_{i=1}^{12} y_i}{36} = \frac{28272 + 28703 + 28523}{36} = 2375 \text{ triệu đồng}$$

Tiếp đến, tính các chỉ số thời vụ cho tháng 1 theo công thức 3.3.12a.

$$I_1 = \frac{\bar{y}_1}{\bar{y}_0} 100 = \frac{1495}{2375} 100 = 62,9 \%$$

Bằng cách tương tự ta tính chỉ số thời vụ cho các tháng còn lại (từ tháng 2 đến tháng 12) trong năm và kết quả được hệ thống ở cột 6 bảng trên.

* Trường hợp biến động thời vụ của các tháng tương ứng giữa các năm có sự tăng (hoặc giảm) rõ rệt thì chỉ số thời vụ được tính theo công thức sau đây:

$$I_i = \frac{\sum_{j=1}^n y_{ij} / \bar{y}_{ij}}{n} 100 \quad ; \quad (3.3.12b)$$

Trong đó:

y_{ij} - Mức độ thực tế ở thời gian i của năm j ;

$\bar{y}_{ij}$ - Mức độ tính toán (có thể là số trung bình trượt hoặc dựa vào phương trình toán học ở thời gian i của năm thứ j);

n - Số năm nghiên cứu.

Hiện nay, với sự trợ giúp của các phần mềm như: SPSS, STATA, Eview, Excel... chúng ta dễ dàng tính toán các tham số trên.

3.4. PHƯƠNG PHÁP PHÂN TÍCH TƯƠNG QUAN

3.4.1. Liên hệ tương quan và phương pháp phân tích tương quan

Mối liên hệ ràng buộc lẫn nhau giữa các chỉ tiêu hoặc tiêu thức của hiện tượng (từ đây chỉ dùng từ "chỉ tiêu" đặc trưng cho cả hai), trong đó sự biến động của một chỉ tiêu này (chỉ tiêu kết quả) là do tác động của nhiều chỉ tiêu khác (các chỉ tiêu nguyên nhân) gọi là liên hệ tương quan - một hình thức liên hệ không chặt chẽ.

Ví dụ: Năng suất lúa tăng lên là do tác động của nhiều nhân tố: Phân bón, giống lúa, làm đất, chăm bón,... thì liên hệ giữa năng suất lúa và các nhân tố nêu trên là quan hệ tương quan; trong đó năng suất lúa là chỉ tiêu kết quả, còn phân bón, giống lúa, chi phí chăm bón, làm đất là các chỉ tiêu nguyên nhân.

Chú ý rằng trong quan hệ tương quan, tác động của các chỉ tiêu nguyên nhân đối với chỉ tiêu kết quả có các mức độ khác nhau: Có chỉ tiêu nguyên nhân gây ảnh hưởng nhiều (tương quan mạnh), có chỉ tiêu nguyên nhân gây ảnh hưởng không đáng kể (tương quan yếu). Điều này phụ thuộc vào tính chất quan hệ của các chỉ tiêu và điều kiện cụ thể của từng trường hợp.

Mục đích cuối cùng của phân tích thống kê là nghiên cứu mối quan hệ giữa các chỉ tiêu khác nhau và xác định mức độ ảnh hưởng của từng chỉ tiêu cũng như mức độ ảnh hưởng của nhiều chỉ tiêu nguyên nhân đến chỉ tiêu kết quả cụ thể như thế nào?

Một phương pháp toán học áp dụng vào việc phân tích thống kê nhằm biểu hiện và nghiên cứu mối liên hệ tương quan giữa các chỉ tiêu của hiện tượng kinh tế - xã hội là phương pháp phân tích tương quan.

Khi phân tích tương quan không thể xác định quan hệ và mức độ

ảnh hưởng lẫn nhau của tất cả các chỉ tiêu của hiện tượng mà chỉ thể hiện trên hai hay một số chỉ tiêu nào đó được xem là chủ yếu (có tương quan mạnh hơn) với giả thiết các chỉ tiêu khác còn lại coi như không thay đổi.

Quá trình phân tích tương quan gồm các công việc cụ thể sau:

- Phân tích định tính về bản chất của mối quan hệ, đồng thời dùng phương pháp phân tổ hoặc đồ thị để xác định mức độ thực tế của mối quan hệ tương quan, tính chất và xu thế của mối quan hệ đó.

- Biểu hiện cụ thể mối liên hệ tương quan bằng một phương trình hồi quy tuyến tính (đường thẳng) hoặc phương trình hồi quy phi tuyến tính (đường cong) và tính các tham số của các phương trình hồi quy nói trên.

- Đánh giá mức độ chặt chẽ của mối liên hệ tương quan bằng các hệ số tương quan hoặc tỉ số tương quan.

Phương pháp tương quan cho phép đánh giá mức độ quan hệ bằng số liệu cụ thể giữa các chỉ tiêu của hiện tượng nghiên cứu. Đây là ưu điểm nổi bật của phương pháp phân tích tương quan, nên phương pháp có thể áp dụng rất rộng rãi và có hiệu quả trong phân tích thống kê kinh tế.

3.4.2. Phân tích mối liên hệ tương quan giữa các tiêu thức biến đổi theo không gian

Liên hệ tương quan giữa các chỉ tiêu biến đổi theo không gian, nghĩa là mối liên hệ của các chỉ tiêu được nghiên cứu trên góc độ các không gian khác nhau và được sắp xếp theo một thứ tự nào đó. Ví dụ, nghiên cứu mối liên hệ giữa tuổi nghề của công nhân với năng suất lao động của họ.

Với liên hệ tương quan không gian, có 3 trường hợp nghiên cứu: Liên hệ tương quan tuyến tính giữa hai chỉ tiêu, liên hệ tương quan phi tuyến tính giữa hai chỉ tiêu và liên hệ tương quan tuyến tính giữa nhiều chỉ tiêu.

3.4.2.1. Liên hệ tương quan tuyến tính giữa hai chỉ tiêu

a. Phương trình hồi quy tuyến tính (đường thẳng)

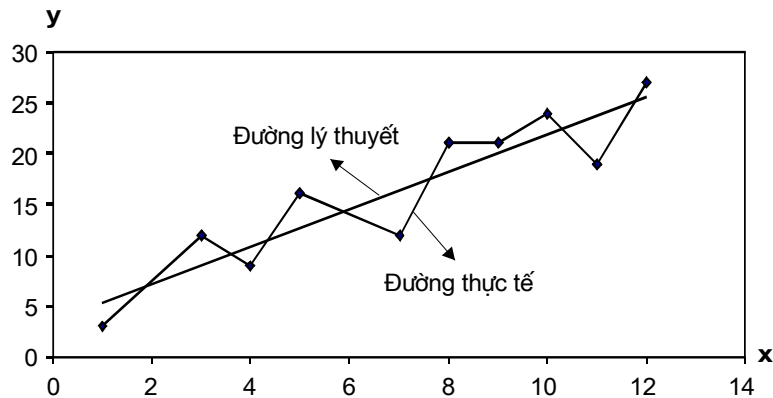
Nếu gọi y và x là các trị số thực tế của chỉ tiêu kết quả và nguyên nhân có thể xây dựng được phương trình hồi quy đường thẳng như sau:

$$\tilde{y}_x = a + bx \quad ; \quad (3.4.1a)$$

Trong đó: $\tilde{y}_x$ là trị số lý thuyết (điều chỉnh) của chỉ tiêu kết quả; a và b là các hệ số của phương trình (trong đó $b > 0$ thì đường thẳng đi lên, $b < 0$ thì đường thẳng đi xuống và $b = 0$ đường thẳng song song với trục hoành).

Có thể biểu diễn giá trị thực tế và giá trị lý thuyết của chỉ tiêu kết quả (qua trục tung) trong quan hệ với chỉ tiêu nguyên nhân (qua trục hoành) qua đồ thị 3.4.1:

Đồ thị 3.4.1: Đặc trưng mối quan hệ giữa chỉ tiêu kết quả (y) và chỉ tiêu nguyên nhân (x)



Bằng phương pháp bình phương nhỏ nhất xây dựng được hệ phương trình chuẩn tắc xác định các hệ số a và b của phương trình đường thẳng như sau:

$$\begin{cases} na + b\Sigma x = \Sigma y \\ a\Sigma x + b\Sigma x^2 = \Sigma xy \end{cases} \quad ; \quad (3.4.1b)$$

Ví dụ: Có số liệu về tuổi nghề và năng suất lao động của các công nhân như cột 1 và 2 bảng 3.4.1:

Bảng 3.4.1. Bảng tính toán các hệ số của phương trình đường thẳng

STT công nhân	Tuổi nghề x (Năm)	Năng suất lao động - y (Triệu đồng)	xy	x^2	y^2
A	1	2	3=1x2	4=(1) ²	5=(2) ²
A	1	3	3	1	9
B	3	12	36	9	144
C	4	9	36	16	81
D	5	16	80	25	256
E	7	12	84	49	144
F	8	21	168	64	441
G	9	21	189	81	441
H	10	24	240	100	576
I	11	19	209	121	361
K	12	27	324	144	729
Tổng	70	164	1369	610	3182

Từ số liệu đã cho của x và y ở bảng 3.4.1, ta tính toán các đại lượng xy, x^2 và y^2 như cột 3, 4 và 5 của bảng.

Thay số liệu tính được ở bảng 3.4.1 vào hệ phương trình 3.4.1b ta có:

$$\begin{cases} 10a + 70b = 164 \\ 70a + 610b = 1369 \end{cases}$$

Giải hệ phương trình tính được: $a = 3,52$ và $b = 1,84$.

Dạng cụ thể của phương trình đường thẳng là:

$$\tilde{y}_x = 3,52 + 1,84x.$$

b. Hệ số tương quan tuyến tính giữa hai chỉ tiêu (ký hiệu là r)

Công thức tính hệ số tương quan:

$$r = \frac{\sum(x - \bar{x})(y - \bar{y})}{\sqrt{\sum(x - \bar{x})^2 \cdot \sum(y - \bar{y})^2}} \quad ; \quad (3.4.2a)$$

Trong đó: $\bar{x} = \frac{\sum x}{n}$ và $\bar{y} = \frac{\sum y}{n}$

Bằng cách biến đổi ta có hệ số tương quan như sau:

$$r = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\delta_x \cdot \delta_y} \quad \text{hoặc} \quad r = b \cdot \frac{\delta_x}{\delta_y} \quad ; \quad (3.4.2b)$$

Trong đó: $\overline{xy} = \frac{\sum xy}{n}$;

$$\delta_x = \sqrt{\frac{\sum(x - \bar{x})^2}{n}} = \sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2} \quad ;$$

$$\delta_y = \sqrt{\frac{\sum(y - \bar{y})^2}{n}} = \sqrt{\frac{\sum y^2}{n} - \left(\frac{\sum y}{n}\right)^2} .$$

Hệ số tương quan có giá trị trong khoảng từ -1 đến 1 ($-1 \leq r \leq 1$):

- Khi r mang dấu dương, giữa x và y có tương quan thuận, khi r mang dấu âm là có tương quan nghịch;

- Khi r càng gần 0 thì quan hệ càng lỏng lẻo, ngược lại khi r càng gần 1 hoặc -1 thì quan hệ càng chặt chẽ. Trường hợp $r = 0$ thì giữa x và y không có quan hệ.

Trở lại ví dụ bảng 3.4.1, ta tính được:

$$\bar{x} = \frac{70}{10} = 7; \quad \bar{y} = \frac{164}{10} = 16,4; \quad \overline{xy} = \frac{1369}{10} = 136,9;$$

$$\delta_x = \sqrt{\frac{610}{10} - \left(\frac{70}{10}\right)^2} = 3,464 \quad \text{và} \quad \delta_y = \sqrt{\frac{3182}{10} - \left(\frac{164}{10}\right)^2} = 7,017$$

Từ số liệu tính toán tiếp tục tính hệ số tương quan (theo công thức 3.4.2b):

$$r = \frac{136,9 - (7 \times 16,4)}{3,464 \times 7,017} = 0,909$$

Theo kết quả tính toán có $r = 0,909$, chứng tỏ giữa tuổi nghề và năng suất lao động của công nhân có mối liên hệ thuận khá chặt chẽ.

3.4.2.2. Liên hệ tương quan phi tuyến tính giữa hai chỉ tiêu

Mối liên hệ tương quan phi tuyến tính, tức là có phương trình hồi quy là đường cong, ví dụ như mối liên hệ giá thành đơn vị sản phẩm và khối lượng sản phẩm: Sự tăng lên của khối lượng sản phẩm có thể dẫn đến việc giảm giá thành đơn vị sản phẩm, nhưng việc giảm này không theo một tỷ lệ tương ứng với sự tăng lên của khối lượng sản phẩm, mà giảm theo tỷ lệ nhỏ dần. Nếu biểu diễn quan hệ giữa 2 chỉ tiêu này lên đồ thị sẽ có dạng hypecbol.

a. Một số phương trình hồi quy phi tuyến

Trong thực tế tùy theo đặc điểm và tính chất của mối quan hệ ta lựa chọn phương trình hồi quy phi tuyến tính phù hợp. Sau đây là một số phương trình hồi quy phi tuyến tính thường được sử dụng:

* Phương trình parabol bậc 2:

$$\tilde{y}_x = a + bx + cx^2 \quad ; \quad (3.4.3a)$$

Phương trình parabol bậc 2 thường được sử dụng khi các trị số của chỉ tiêu nguyên nhân tăng lên thì trị số của chỉ tiêu kết quả tăng (hoặc giảm), việc tăng (hoặc giảm) đạt đến trị số cực đại (hoặc cực tiểu) rồi sau đó lại giảm (hoặc tăng). Ví dụ, nghiên cứu mối liên hệ giữa lượng tiêu hao than và chất lượng gạch máy. Khi lượng tiêu hao than cho 1000 viên gạch còn thấp thì nếu tăng lượng tiêu hao than sẽ làm cho gạch nung ra già hơn, chất lượng cao hơn. Nhưng tăng lượng tiêu hao than đạt đến một mức nào đó (vừa đủ), nếu tiếp tục tăng nữa thì sẽ làm cho gạch nung ra bị khô phồng tức là làm cho chất lượng gạch lại giảm đi. Khi lượng tiêu hao than đạt đến mức vừa đủ thì gạch máy sẽ đạt chất lượng cao nhất (đạt giá trị cực đại).

Bằng phương pháp bình phương nhỏ nhất ta xây dựng được hệ phương trình chuẩn tắc để xác định các hệ số a, b và c của phương trình hồi quy 3.4.3a như sau:

$$\begin{cases} na + b\sum x + c\sum x^2 &= \sum y \\ a\sum x + b\sum x^2 + c\sum x^3 &= \sum xy \\ a\sum x^2 + b\sum x^3 + c\sum x^4 &= \sum x^2 y \end{cases} \quad ; \quad (3.4.3b)$$

* *Phương trình hypecbol:*

$$\tilde{y}_x = a + \frac{b}{x} \quad ; \quad (3.4.4a)$$

Phương trình hypecbol được áp dụng trong trường hợp khi các trị số của chỉ tiêu nguyên nhân tăng lên thì trị số của chỉ tiêu kết quả giảm nhưng mức độ giảm nhỏ dần và đến một giới hạn nào đó ($\tilde{y}_x = a$) thì hầu như không giảm. Ví dụ, quan hệ giữa giá thành đơn vị sản phẩm và khối lượng sản phẩm sản xuất là quan hệ theo phương trình hypecbol như đã nói ở trên.

Bằng phương pháp bình phương nhỏ nhất ta xây dựng được hệ phương trình chuẩn tắc xác định hệ số a và b của phương trình (3.4.4a) như sau:

$$\begin{cases} na + b\sum \frac{1}{x} &= \sum y \\ a\sum \frac{1}{x} + b\sum \frac{1}{x^2} &= \sum \frac{y}{x} \end{cases} \quad ; \quad (3.4.4b)$$

* *Phương trình hàm số mũ:*

$$\tilde{y}_x = a.b^x \quad ; \quad (3.4.5a)$$

Phương trình hàm số mũ được áp dụng trong trường hợp cùng với sự tăng lên của chỉ tiêu nguyên nhân thì trị số của các chỉ tiêu kết quả thay đổi theo cấp số nhân, nghĩa là có tốc độ tăng xấp xỉ nhau.

Bằng phương pháp bình phương nhỏ nhất xây dựng được hệ phương trình chuẩn tắc để xác định các hệ số của phương trình hồi quy như sau:

$$\begin{cases} n \ln a + \ln b \sum x = \sum \ln y \\ \ln a \sum x + \ln b \sum x^2 = \sum x \ln y \end{cases} \quad ; \quad (3.4.5b)$$

b. Tỉ số tương quan

Đối với liên hệ tương quan phi tuyến tính giữa 2 chỉ tiêu sẽ dùng tỉ số tương quan (ký hiệu $\eta = \text{eta}$) để đánh giá trình độ chặt chẽ của mối liên hệ. Công thức tính tỉ số tương quan như sau:

$$\eta = \sqrt{\frac{\delta_{\tilde{y}_x}^2}{\delta_y^2}} = \frac{\delta_{\tilde{y}_x}}{\delta_y} \quad ; \quad (3.4.6)$$

Trong đó:

$$- \delta_{\tilde{y}_x}^2 = \frac{(\tilde{y}_x - \bar{y})^2}{n} : \text{Phương sai đo độ biến thiên của chỉ tiêu } y \text{ do}$$

ảnh hưởng riêng của chỉ tiêu x; với $\tilde{y}_x$ là giá trị lý thuyết của đường hồi quy phi tuyến tính giữa y và x được xác định;

$$- \delta_y^2 = \frac{(y - \bar{y})^2}{n} : \text{Phương sai đo độ biến thiên của chỉ tiêu } y \text{ do}$$

ảnh hưởng của tất cả các chỉ tiêu nguyên nhân.

Tỉ số tương quan có một số tính chất sau:

+ Tỉ số tương quan lấy giá trị trong khoảng $[0;1]$, tức là $0 \leq \eta \leq 1$.

- Nếu $\eta = 0$ thì giữa x và y không có liên hệ tương quan;

- Nếu $\eta = 1$ thì giữa x và y có liên hệ hàm số;

- Nếu η càng gần 1 thì giữa x và y có liên hệ tương quan càng chặt chẽ và càng gần 0 thì liên hệ tương quan càng lỏng lẻo.

+ Tỉ số tương quan lớn hơn hoặc bằng giá trị tuyệt đối của hệ số tương quan, tức là $\eta \geq |r|$. Nếu $\eta = |r|$ thì giữa x và y có mối liên hệ tương quan tuyến tính.

3.4.2.3. Liên hệ tương quan tuyến tính giữa nhiều chỉ tiêu

Trong thực tế các hiện tượng kinh tế - xã hội, một chỉ tiêu kết quả thường do tác động của nhiều chỉ tiêu nguyên nhân. Ví dụ, năng suất lao động của công nhân tăng lên do ảnh hưởng của các yếu tố nguyên nhân: Tuổi nghề, trình độ trang bị kỹ thuật, trình độ quản lý, v.v... Do đó vấn đề đặt ra là cần phải nghiên cứu mối liên hệ giữa một chỉ tiêu kết quả với một số chỉ tiêu nguyên nhân.

Để dễ theo dõi, dưới đây chỉ trình bày nội dung và phương pháp phân tích mối liên hệ tương quan giữa ba chỉ tiêu.

a. Phương trình hồi quy tuyến tính giữa ba chỉ tiêu

Nếu gọi y là chỉ tiêu kết quả và x_1, x_2 là các chỉ tiêu nguyên nhân, ta có phương trình hồi quy tuyến tính giữa 3 chỉ tiêu như sau:

$$\tilde{y}_{x_1, x_2} = a_0 + a_1 \cdot x_1 + a_2 \cdot x_2 \quad ; \quad (3.4.7a)$$

Bằng phương pháp bình phương nhỏ nhất, xây dựng được hệ phương trình chuẩn tắc để tính các tham số của phương trình hồi quy 3.4.7a như sau:

$$\begin{cases} na_0 + a_1 \Sigma x_1 + a_2 \Sigma x_2 = \Sigma y \\ a_0 \Sigma x_1 + a_1 \Sigma x_1^2 + a_2 \Sigma x_1 x_2 = \Sigma x_1 y \\ a_0 \Sigma x_2 + a_1 \Sigma x_1 x_2 + a_2 \Sigma x_2^2 = \Sigma x_2 y \end{cases} \quad ; \quad (3.4.7b)$$

b. Hệ số tương quan

Để đánh giá trình độ chặt chẽ của mối liên hệ tương quan tuyến tính nhiều chỉ tiêu, người ta thường tính toán các hệ số tương quan gồm: Hệ số tương quan bội và hệ số tương quan riêng.

* *Hệ số tương quan bội* (Ký hiệu là R) được dùng để đánh giá trình độ chặt chẽ giữa chỉ tiêu kết quả với tất cả các chỉ tiêu nguyên nhân được nghiên cứu. Công thức tính như sau:

$$R = \sqrt{\frac{r_{yx_1}^2 + r_{yx_2}^2 - 2r_{yx_1} r_{yx_2} r_{x_1 x_2}}{1 - r_{x_1 x_2}^2}} \quad ; \quad (3.4.8)$$

Trong đó: r_{yx_1}, r_{yx_2} và $r_{x_1 x_2}$ là các hệ số tương quan tuyến tính giữa các cặp tiêu thức y với x_1 , y với x_2 và x_1 với x_2 (tính như các công thức 3.4.2a, 3.4.2b hoặc 3.4.2c).

Hệ số tương quan bội nhận giá trị trong khoảng $[0;1]$, tức là $0 \leq R \leq 1$.

Như vậy, R càng gần 0 thì quan hệ tương quan càng lỏng lẻo và R càng gần 1 thì quan hệ càng chặt chẽ.

Nếu $R = 0$ thì không có quan hệ tương quan và nếu $R = 1$ thì quan hệ tương quan trở thành quan hệ hàm số.

* *Hệ số tương quan riêng* được dùng để đánh giá trình độ chặt chẽ của mối liên hệ giữa tiêu thức kết quả với từng tiêu thức nguyên nhân trong điều kiện đã loại trừ ảnh hưởng của các tiêu thức nguyên nhân khác. Trong trường hợp mối liên hệ giữa y với x_1 và x_2 ở trên có

thể tính:

- Hệ số tương quan riêng giữa y và x_1 (loại trừ ảnh hưởng của x_2):

$$r_{yx_1(x_2)} = \frac{r_{yx_1} - r_{yx_2} \times r_{x_1x_2}}{\sqrt{(1 - r_{yx_2}^2) \cdot (1 - r_{x_1x_2}^2)}} \quad ; \quad (3.4.9a)$$

- Hệ số tương quan riêng giữa y và x_2 (loại trừ ảnh hưởng của x_1):

$$r_{yx_2(x_1)} = \frac{r_{yx_2} - r_{yx_1} \times r_{x_1x_2}}{\sqrt{(1 - r_{yx_1}^2) \cdot (1 - r_{x_1x_2}^2)}} \quad ; \quad (3.4.9b)$$

Ví dụ: Có tài liệu về năng suất lao động, phần trăm chi phí nguyên vật liệu nhập ngoại trong giá thành sản phẩm và giá thành đơn vị sản phẩm của 5 doanh nghiệp cùng sản xuất ra 1 loại sản phẩm như bảng 3.4.2:

Bảng 3.4.2: Một số chỉ tiêu của 5 doanh nghiệp

Thứ tự Doanh nghiệp	Năng suất lao động (x_1 - Triệu đồng)	% nguyên vật liệu nhập ngoại - x_2 (%)	Giá thành đơn vị (y - Nghìn đồng)
1	20	52	44
2	21	51	43
3	23	51	42
4	25	50	40
5	26	51	41
Tổng số	115	255	210
Số bình quân	23	51	42
Độ lệch chuẩn	2,28	0,63	1,41

Từ số liệu đã cho ở bảng 3.4.2 ta lập bảng tính toán 3.4.3:

Bảng 3.4.3: Bảng tính các đại lượng cho hệ phương trình

TT Doanh nghiệp	x_1y	x_2y	x_1x_2	x_1^2	x_2^2	y^2
1	880	2288	1040	400	2704	1936
2	903	2193	1071	441	2601	1849
3	966	2142	1173	529	2601	1764
4	1000	2000	1250	625	2500	1600
5	1066	2091	1326	676	2601	1681
Tổng số	4815	10714	5860	2671	13007	8830
Số BQ	963	2142,8	1172	534,2	2691,4	1766

Thay số liệu vào hệ phương trình chuẩn tắc 3.4.7b ta có:

$$\begin{cases} 5a_0 + 115a_1 + 255a_2 = 210 \\ 115a_0 + 2671a_1 + 5860a_2 = 4815 \\ 255a_0 + 5860a_1 + 13007a_2 = 10714 \end{cases}$$

Giải hệ phương trình ta được: $a_0 = -4,26$; $a_1 = -0,37$; $a_2 = 1,07$.

Do đó:

Phương trình hồi quy:

$$\tilde{y}_{x_1x_2} = -4,26 - 0,37x_1 + 1,07x_2$$

Các hệ số tương quan:

- Các hệ số tương quan tuyến tính giữa hai tiêu thức:

$$r_{yx_1} = \frac{\overline{x_1y} - \overline{x_1} \cdot \overline{y}}{\sigma_{x_1} \cdot \sigma_y} = \frac{963 - 23 \times 42}{2,28 \times 1,41} = -0,94$$

$$r_{yx_2} = \frac{\overline{x_2 y} - \overline{x_2} \cdot \overline{y}}{\sigma_{x_2} \cdot \sigma_y} = \frac{2142,8 - 51 \times 42}{0,63 \times 1,41} = 0,89$$

$$r_{x_1 x_2} = \frac{\overline{x_1 x_2} - \overline{x_1} \cdot \overline{x_2}}{\sigma_{x_1} \cdot \sigma_{x_2}} = \frac{1172 - 23 \times 51}{2,28 \times 0,63} = -0,69$$

- Hệ số tương quan bội:

$$R_{yx_1 x_2} = \sqrt{\frac{(-0,94)^2 + (0,89)^2 - 2 \times (-0,94) \times 0,89 \times (-0,69)}{1 - (-0,69)^2}} = 0,985$$

- Các hệ số tương quan riêng:

$$r_{yx_1(x_2)} = \frac{-0,94 - 0,89 \times (-0,69)}{\sqrt{(1 - 0,89^2)(1 - 0,69^2)}} = -0,987$$

$$r_{yx_2(x_1)} = \frac{0,89 - (-0,94) \times (-0,69)}{\sqrt{(1 - 0,94^2)(1 - 0,69^2)}} = 0,88$$

Các kết quả tính toán ở trên cho thấy mối liên hệ giữa giá thành đơn vị sản phẩm với năng suất lao động và tỷ lệ phần trăm nguyên, vật liệu nhập ngoại trong giá thành rất chặt chẽ ($R_{yx_1 x_2} = 0,985$). Trong mối liên hệ này thì năng suất lao động tỷ lệ nghịch với giá thành đơn vị sản phẩm, còn tỷ lệ giá trị nguyên, vật liệu nhập ngoại tỷ lệ thuận với giá thành đơn vị sản phẩm.

3.4.3. Phân tích mối liên hệ tương quan giữa hai chỉ tiêu biến động theo thời gian

Mối liên hệ tương quan theo thời gian là mối liên hệ không chặt chẽ giữa các dãy số biến động theo thời gian; trong đó có một số dãy số biểu hiện biến động của các chỉ tiêu nguyên nhân (sự biến động của nó sẽ ảnh hưởng đến biến động của chỉ tiêu kết quả) và một dãy số biểu hiện biến động của chỉ tiêu kết quả (sự biến động của nó phụ

thuộc vào biến động của các chỉ tiêu nguyên nhân).

Phân tích mối liên hệ tương quan giữa các dãy số theo thời gian chính là xác định mức độ chặt chẽ của mối liên hệ giữa các dãy số. Do đặc điểm nghiên cứu tương quan theo dãy số thời gian là rất phức tạp nên ở đây chỉ trình bày tương quan tuyến tính giữa hai dãy số.

Đặc điểm của dãy số biến động theo thời gian là tồn tại hiện tượng tự tương quan giữa các mức độ của dãy số. Để kiểm tra hiện tượng này ta tiến hành tính hệ số tương quan tuyến tính giữa các mức độ của dãy số đã cho (x_t hoặc y_t) với mức độ của dãy số đó nhưng lệch đi thời gian 1 năm ($t = 1$). Khi nghiên cứu riêng cho từng dãy (đại lượng x hay y) về bản chất đều có công thức tính giống nhau, chỉ khác nhau là theo x hoặc theo y . Từ đây các trường hợp nghiên cứu riêng của từng dãy thống nhất chỉ ký hiệu chung là x .

Công thức tính hệ số tự tương quan riêng cho từng dãy số chẳng hạn x như sau:

$$r_{x_t, x_{t+1}} = \frac{\overline{x_t \cdot x_{t+1}} - \overline{x_t} \cdot \overline{x_{t+1}}}{\sigma_t \cdot \sigma_{t+1}} ; \quad (3.4.10)$$

Trong đó:

t - Chỉ thứ tự thời gian theo từng năm;

x_t, x_{t+1} - Mức độ thực tế của dãy thuộc năm t và của năm sau năm t ($t+1$);

σ_t và σ_{t+1} - Các độ lệch chuẩn tương ứng;

$r_{x_t, x_{t+1}}$ - Hệ số phản ánh mức độ tự tương quan.

Trị số của hệ số này càng gần 1 thì đặc điểm tự tương quan càng mạnh và ngược lại càng gần 0 thì đặc điểm tự tương quan càng yếu.

Khi kiểm tra đặc điểm tự tương quan của dãy số ta xét hai khả năng:

* Nếu thấy đặc điểm này yếu ($r_{x_t, x_{t+1}}$ gần 0) thì hệ số tương

quan tuyến tính giữa hai dãy x_t và y_t ($r_{x,y}$) vẫn tính trực tiếp theo các mức độ thực tế (x_t và y_t) như tương quan tuyến tính giữa hai chỉ tiêu biến động theo không gian đã trình bày ở trên.

$$r_{xy} = \frac{\overline{x.y} - \bar{x}.\bar{y}}{\sigma_x \cdot \sigma_y} \quad ; \quad (3.4.11)$$

Trong đó các đại lượng được tính như sau:

$$\overline{x.y} - \text{Trung bình của tích } x \text{ và } y; \quad \overline{x.y} = \frac{\sum x.y}{n}$$

$$\bar{x} - \text{Trung bình của } x; \quad \bar{x} = \frac{\sum x}{n}$$

$$\bar{y} - \text{Trung bình của } y; \quad \bar{y} = \frac{\sum y}{n}$$

σ_x - Độ lệch chuẩn của các mức độ riêng biệt với mức độ bình quân chung của x .

$$\sigma_x = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}}$$

σ_y - Độ lệch chuẩn của các mức độ riêng biệt với mức độ bình quân chung của y .

$$\sigma_y = \sqrt{\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n}}$$

* Nếu thấy đặc điểm tự tương quan của hai dãy số mạnh ($r_{x_t, x_{t+1}}$ gần 1 hoặc -1) thì hệ số tương quan giữa hai dãy x_t và y_t không thể tính trực tiếp theo các mức độ thực tế (x_t và y_t) mà theo các độ lệch giữa mức độ thực tế (x_t, y_t) và mức độ lý thuyết tương ứng ($\hat{x}_t, \hat{y}_t$). Công thức tính như sau:

$$R_{xy} = \frac{\sum d_{x_t} \cdot d_{y_t}}{\sqrt{\sum d_{x_t}^2 \cdot \sum d_{y_t}^2}} \quad ; \quad (3.4.12)$$

Trong đó: d_{x_t}, d_{y_t} là các độ lệch giữa mức độ thực tế (x_t, y_t) và các mức độ lý thuyết tương ứng ($\hat{x}_t, \hat{y}_t$), tức là $d_{x_t} = x_t - \hat{x}_t$ và $d_{y_t} = y_t - \hat{y}_t$.

Các mức độ lý thuyết $\hat{x}_t$ và $\hat{y}_t$ có thể xác định được bằng nhiều phương pháp, nhưng phổ biến và có ý nghĩa nhất là bằng phương pháp điều chỉnh dãy số theo phương trình toán học (phương trình hồi quy).

Trong kinh tế thường dùng một số dạng, phương trình toán học chủ yếu sau đây để điều chỉnh các dãy số:

$$\text{- Phương trình tuyến tính: } \hat{y} = a_0 + a_1 t \quad ; \quad (3.4.13a)$$

- Phương trình parabol bậc hai:

$$\hat{y} = a_0 + a_1 t + a_2 t^2 \quad ; \quad (3.4.13b)$$

- Phương trình parabol bậc ba:

$$\hat{y} = a_0 + a_1 t + a_2 t^2 + a_3 t^3 \quad ; \quad (3.4.13c)$$

$$\text{- Phương trình hypecbol: } \hat{y} = a_0 + \frac{a_1}{t} \quad ; \quad (3.4.13d)$$

$$\text{- Phương trình hàm số mũ: } \hat{y} = a_0 \cdot a_1^t \quad ; \quad (3.4.13e)$$

Phương pháp tính các hệ số theo từng dạng phương trình trên đã được trình bày ở điểm 3.3.3 mục 3.3 của phần này.

Để xác định quy luật phát triển của từng dãy số theo loại phương trình này, trước tiên phải đưa số liệu lên đồ thị. Nếu quan sát trên dãy số phát triển rõ nét theo một loại phương trình nào đó thì có thể điều chỉnh dãy số một lần. Trường hợp khó xác định một cách cụ thể theo

một loại phương trình nào đó thì phải tiến hành điều chỉnh dãy số theo một số phương trình. Sau đó ứng với mỗi phương trình đã được điều chỉnh chúng ta tính toán các sai số mô tả:

$V_x = \frac{\sigma_x}{\bar{X}}$ và $V_y = \frac{\sigma_y}{\bar{Y}}$ rồi chọn phương trình nào có hệ số mô tả nhỏ nhất.

Dưới đây là ví dụ tính toán hệ số tương quan tuyến tính phản ánh mối liên hệ giữa hai dãy số biến động theo thời gian: mức trang bị vốn cho người lao động và năng suất lao động của công nghiệp Việt Nam từ 1990 đến 2003.

Bảng 3.4.4: Mức trang bị vốn và năng suất lao động của công nghiệp VN

Đơn vị: Triệu đồng

Năm	Thứ tự năm t	Mức trang bị vốn x_i	Năng suất lao động y_i	Năm	Thứ tự năm t	Mức trang bị vốn x_i	Năng suất lao động y_i
A	B	1	2	A	B	1	2
1990	1	25,18	12,97	1997	8	58,97	28,65
1991	2	30,96	15,61	1998	9	64,30	29,96
1992	3	35,44	18,71	1999	10	69,72	30,40
1993	4	41,33	21,69	2000	11	75,30	32,60
1994	5	46,37	24,50	2001	12	83,35	35,21
1995	6	50,45	25,78	2002	13	85,14	35,58
1996	7	53,75	26,84	2003	14	87,28	36,45

Từ số liệu bảng 3.4.4 ta lần lượt tính theo các bước sau:

Bước 1. Kiểm tra tính chất tự tương quan của 2 dãy số trên.

Áp dụng công thức 3.4.10 ta tính được các hệ số tự tương quan:

Dãy x_t : $R_{x_t, x_{t+1}} = 0,9965$ và dãy y_t : $R_{y_t, y_{t+1}} = 0,9942$

Kết quả tính toán trên chứng tỏ cả 2 dãy số đều có tính chất tự tương quan rất mạnh.

Bước 2. Tiến hành hồi quy hai dãy số về mức năng suất lao động và mức trang bị vốn cho lao động theo các dạng hàm: Tuyến tính, hàm bậc hai và hàm số mũ. Kết quả tính toán cho thấy cả hai dãy số năng suất lao động và mức trang bị vốn của lao động hồi quy theo hàm parabol bậc hai có hệ số mô tả nhỏ nhất, tức là có hệ số xác định lớn nhất.

Vậy hàm số được lựa chọn để điều chỉnh biến động của hai dãy số như sau:

- Đối với dãy số x_t :

$\hat{x}_t = 20,6536 + 4,9791 t + 0,0044 t^2$; (3.4.14a)

- Đối với dãy y_t :

$\hat{y}_t = 10,71973 + 2,86166 t - 0,0745 t^2$; (3.4.14b)

Bước 3. Từ các dạng hàm lý thuyết 3.4.14a và 3.4.14b, lần lượt thay giá trị t từ 1 đến 13 vào tính được các giá trị lý thuyết về mức trang bị vốn ($\hat{x}_t$) và năng suất lao động ($\hat{y}_t$) như số liệu cột 3 và 4 bảng 3.4.5.

Bảng 3.4.5: Độ lệch giữa giá trị thực tế và lý thuyết của mức trang bị vốn và năng suất lao động

Đơn vị tính: Triệu đồng

Năm	Giá trị thực tế		Giá trị lý thuyết		Độ lệch giữa thực tế và lý thuyết	
	Mức trang bị vốn x_i	Năng suất lao động y_i	Mức trang bị vốn $\hat{x}_i$	Năng suất lao động $\hat{y}_i$	Mức trang bị vốn d_{x_i}	Năng suất lao động d_{y_i}

A	1	2	3	4	5	6
1990	25,18	12,97	25,6284	13,5069	-0,4460	-0,5391
1991	30,96	15,61	30,5944	16,1450	0,3668	-0,5318
1992	35,44	18,71	35,5517	18,6342	-0,1164	0,0718
1993	41,33	21,69	40,5003	20,9744	0,8344	0,7203
1994	46,37	24,50	45,4402	23,1655	0,9268	1,3301
1995	50,45	25,78	50,3714	25,2077	0,0802	0,5701
1996	53,75	26,84	55,2938	27,1009	-1,5480	-0,2574
1997	58,97	28,65	60,2076	28,8450	-1,2368	-0,1996
1998	64,30	29,96	65,1126	30,4402	-0,8163	-0,4850
1999	69,72	30,40	70,0089	31,8864	-0,2882	-1,4899
2000	75,30	32,60	74,8965	33,1835	0,4010	-0,5811
2001	83,35	35,21	79,7754	34,3317	3,5736	0,8736
2002	85,14	35,58	84,6456	35,3309	0,4912	0,2454
2003	87,28	36,45	89,5071	36,1810	-2,2223	0,2725

Từ số liệu theo giá trị thực tế và giá trị lý thuyết của mức trang bị vốn và năng suất lao động ta tính được các độ lệch tương ứng ở cột 5 và 6 bảng 3.4.5.

Bước 4. Tính hệ số tương quan giữa năng suất lao động và mức trang bị vốn.

Từ số liệu bảng 3.4.5 về các giá trị d_{xi} và d_{yi} ta tiếp tục lập bảng xác định các đại lượng để tính hệ số tương quan.

Bảng 3.4.6: Xác định các đại lượng để tính hệ số tương quan

STT	d_{xi}	d_{yi}	d_{xi}^2	d_{yi}^2	$d_{xi} \cdot d_{yi}$
1	-0,4460	-0,5391	0,1989	0,2907	0,2405
2	0,3668	-0,5318	0,1345	0,2828	-0,1950
3	-0,1164	0,0718	0,0135	0,0051	-0,0083

4	0,8344	0,7203	0,6962	0,5189	0,6010
5	0,9268	1,3301	0,8590	1,7692	1,2328
6	0,0802	0,5701	0,0064	0,3250	0,0457
7	-1,5480	-0,2574	2,3965	0,0662	0,3984
8	-1,2368	-0,1996	1,5297	0,0398	0,2468
9	-0,8163	-0,4850	0,6663	0,2352	0,3959
10	-0,2882	-1,4899	0,0831	2,2197	0,4294
11	0,4010	-0,5811	0,1608	0,3377	-0,2330
12	3,5736	0,8736	12,7707	0,7632	3,1219
13	0,4912	0,2454	0,2412	0,0602	0,1205
14	-2,2223	0,2725	4,9384	0,0743	-0,6057
Tổng cộng	x	x	24,6953	6,9879	5,7909

Theo số liệu bảng 3.4.6, áp dụng công thức 3.4.12 ta tính được hệ số tương quan:

$$R_{xy} = \frac{5,7909}{\sqrt{24,6953 \cdot 6,9879}} = 0,4408$$

Hệ số tương quan bằng 0,4408 chứng tỏ mối quan hệ giữa năng suất lao động và mức trang bị vốn cố định cho lao động của ngành công nghiệp tương đối chặt chẽ.

3.5. PHƯƠNG PHÁP CHỈ SỐ

3.5.1. Một số vấn đề chung về phương pháp chỉ số

Chỉ số trong thống kê là chỉ tiêu tương đối biểu hiện quan hệ so sánh giữa các mức độ của một hiện tượng kinh tế - xã hội. Chỉ số tính được bằng cách so sánh hai mức độ của hiện tượng ở hai thời gian hoặc không gian khác nhau, nhằm nêu lên sự biến động của hiện tượng qua thời gian hoặc không gian.

*** Ý nghĩa của chỉ số trong thống kê**

- Nghiên cứu sự biến động về mức độ của hiện tượng qua thời gian (biến động của giá cả, giá thành, năng suất lao động, khối lượng sản phẩm, diện tích gieo trồng,...). Các chỉ số tính theo mục đích này thường gọi là chỉ số phát triển.

- So sánh chênh lệch về mức độ của hiện tượng qua không gian (chênh lệch giá cả, lượng hàng hoá tiêu thụ giữa hai thị trường, giữa hai địa phương, hai khu vực,...). Các chỉ số tính theo mục đích này thường gọi là chỉ số không gian.

- Xác định nhiệm vụ kế hoạch hoặc đánh giá kết quả thực hiện kế hoạch về các chỉ tiêu kinh tế - xã hội. Các chỉ số này thường gọi là chỉ số kế hoạch.

- Phân tích mức độ ảnh hưởng và xác định vai trò đóng góp của các nhân tố khác nhau đối với sự biến động chung của hiện tượng phức tạp (ví dụ: Xác định xem sự biến động của các nhân tố năng suất lao động và số lượng công nhân đã ảnh hưởng đến mức độ nào đối với sự tăng giảm của kết quả sản xuất do công nhân tạo ra). Thực chất đây cũng là phân tích mối liên hệ của các yếu tố nguyên nhân với nhau cũng như tính toán ảnh hưởng của mỗi yếu tố nguyên nhân đến chỉ tiêu kết quả.

*** Một số hình thức phân loại chủ yếu về chỉ số**

- Căn cứ theo phạm vi tính toán của chỉ số: Chia thành chỉ số cá thể và chỉ số tổng hợp (xem chỉ số cá thể và chỉ số tổng hợp).

- Căn cứ tính chất của chỉ tiêu cấu thành tổng thể: Chia thành chỉ số chỉ tiêu chất lượng và chỉ số chỉ tiêu khối lượng (việc phân thành chỉ tiêu chất lượng và khối lượng chỉ có ý nghĩa tương đối).

- Căn cứ hình thức biểu hiện, chia thành chỉ số ở dạng cơ bản và chỉ số ở dạng biến đổi (xem chỉ số tổng hợp và chỉ số bình quân).

- Căn cứ thời kỳ gốc so sánh, chia thành chỉ số liên hoàn và chỉ số

định gốc (xem chỉ số liên hoàn và chỉ số định gốc).

- Căn cứ số lượng nhân tố lượng biến của hiện tượng, chia thành chỉ số chung và chỉ số nhân tố (xem hệ thống các chỉ số).

* **Đặc điểm của phương pháp chỉ số** là biểu hiện về lượng của các phần tử trong hiện tượng phức tạp được chuyển về dạng chung có thể trực tiếp cộng được với nhau, dựa trên cơ sở mối quan hệ giữa nhân tố nghiên cứu với các nhân tố khác. Ví dụ: Khối lượng sản phẩm các loại, vốn không thể trực tiếp cộng được với nhau, khi được chuyển sang dạng giá trị, bằng cách nhân với yếu tố giá cả để có thể trực tiếp cộng với nhau. Mặt khác, khi nghiên cứu biến động của một nhân tố, bằng cách giả định các nhân tố khác của hiện tượng phức tạp không thay đổi, nhờ đó phương pháp chỉ số cho phép loại trừ ảnh hưởng biến động của các nhân tố này để khảo sát sự biến động riêng biệt của các nhân tố cần nghiên cứu.

* **Trong công thức chỉ số tổng hợp, nhân tố** biểu hiện sự biến động về mức độ của hiện tượng nghiên cứu gọi là lượng biến của chỉ số. Ví dụ: Trong chỉ số giá cả, lượng biến của chỉ số là giá cả các loại hàng, trong chỉ số khối lượng sản phẩm, lượng biến của chỉ số là khối lượng sản phẩm mỗi loại.

* **Trong công thức chỉ số tổng hợp, nhân tố quan hệ trực tiếp với lượng biến của chỉ số**, được cố định ở một thời kỳ nào đó ở cả tử số và mẫu số của chỉ số gọi là quyền số. Ví dụ: Trong chỉ số giá cả, quyền số là khối lượng hàng hoá tiêu thụ kỳ báo cáo; trong chỉ số khối lượng sản phẩm, quyền số là giá cả kỳ gốc.

Trong một chỉ số, quyền số có thể là một nhân tố (ví dụ, trong chỉ số tổng hợp về giá cả (xem chỉ số tổng hợp), quyền số là lượng hàng hoá tiêu thụ hoặc trong chỉ số tổng hợp về lượng hàng hoá tiêu thụ, quyền số là giá cả (xem chỉ số tổng hợp)); nhưng cũng có thể là tích của nhiều nhân tố khác nhau, (ví dụ, trong chỉ số bình quân điều hoà gia quyền về giá cả, chỉ số bình quân số học về khối lượng sản phẩm,

quyền số đều là tích của giá cả và lượng hàng hoá tiêu thụ (p.q) (xem các chỉ số bình quân)).

Quyền số của chỉ số có thể giải quyết hai nhiệm vụ:

- Chuyển các phần tử vốn không trực tiếp cộng được với nhau thành dạng chung để có thể cộng được với nhau;
- Nói lên tầm quan trọng của mỗi phần tử trong toàn bộ tổng thể.

3.5.2. Chỉ số cá thể và chỉ số tổng hợp

3.5.2.1. Chỉ số cá thể

Chỉ số cá thể là chỉ tiêu tương đối biểu hiện sự biến động của từng phần tử, từng đơn vị cá biệt trong một tổng thể phức tạp.

Ví dụ:

A. Chỉ số giá bán của từng loại mặt hàng:

$$i_p = \frac{p_1}{p_0} \quad ; \quad (3.5.1)$$

Trong đó: p_1, p_0 - Giá bán kỳ báo cáo và kỳ gốc.

B. Chỉ số khối lượng hàng hoá tiêu thụ của từng mặt hàng:

$$i_q = \frac{q_1}{q_0} \quad ; \quad (3.5.2)$$

Trong đó: q_1, q_0 - Lượng hàng hoá tiêu thụ kỳ báo cáo và kỳ gốc.

Chỉ số cá thể cũng được nghiên cứu theo thời gian, không gian và theo kế hoạch.

Thực chất của chỉ số cá thể là các số tương đối động thái (nghiên cứu biến động theo thời gian), số tương đối không gian (nghiên cứu biến động theo không gian) và số tương đối kế hoạch (nghiên cứu biến động của thực tế so với kế hoạch). Do vậy tính toán rất đơn giản và áp dụng thuận tiện.

Hạn chế của chỉ số cá thể là chỉ nghiên cứu biến động riêng của từng phần tử, từng đơn vị cá biệt trong tổng thể, không cho phép ta nghiên cứu biến động chung của nhiều phần tử, hoặc nhiều đơn vị trong một tổng thể gồm các phần tử, hoặc các đơn vị không thể trực tiếp cộng được với nhau để so sánh. Ví dụ, một cửa hàng tiêu thụ 3 loại mặt hàng: Vải (tính bằng mét); dầu gội đầu (tính bằng lọ) và xà phòng (tính bằng kg). Chỉ số cá thể chỉ cho phép tính toán tốc độ phát triển riêng của từng mặt hàng đó, chứ không cho phép cộng trực tiếp 3 mặt hàng đó lại với nhau để so sánh nhằm xác định tốc độ phát triển chung của cả 3 loại mặt hàng này vì chúng có giá trị sử dụng cũng như có đơn vị tính khác nhau.

3.5.2.2. Chỉ số tổng hợp

Chỉ số tổng hợp là chỉ tiêu tương đối phản ánh sự biến động một nhân tố (như ở trên đã nói là lượng biến) của hiện tượng kinh tế - xã hội phức tạp. Các nhân tố khác còn lại được cố định ở một thời kỳ nào đó gọi là quyền số.

Quyền số có thể được chọn ở các kỳ khác nhau (kỳ gốc, kỳ báo cáo, kỳ kế hoạch hoặc một kỳ nào đó thích hợp) tùy theo mục đích nghiên cứu. Thời kỳ của quyền số có ảnh hưởng nhất định đến trị số và khả năng tính toán của chỉ số. Do đó việc chọn thời kỳ của quyền số tùy thuộc vào yêu cầu nghiên cứu và điều kiện về số liệu cụ thể.

Dưới đây sẽ trình bày các công thức tính chỉ số tổng hợp theo các hình thức lựa chọn thời kỳ quyền số khác nhau được bắt đầu từ một ví dụ nghiên cứu hiện tượng có 2 yếu tố: Giá cả và lượng hàng hoá tiêu thụ (trong quan hệ này giá là chỉ tiêu chất lượng, còn lượng hàng hoá tiêu thụ là chỉ tiêu số lượng).

a. Chỉ số tổng hợp về giá cả

*** Chỉ số tổng hợp về giá cả theo thời gian**

- Nếu chọn quyền số là lượng hàng hoá tiêu thụ ở kỳ gốc, chỉ số tổng hợp về giá cả theo Laspeyres có dạng sau:

$$I_p = \frac{\sum p_1 q_0}{\sum p_0 q_0} \quad ; \quad (3.5.3)$$

- Nếu chọn quyền số là lượng hàng hoá tiêu thụ kỳ báo cáo, chỉ số tổng hợp về giá cả theo Paashe có dạng sau:

$$I_p = \frac{\sum p_1 q_1}{\sum p_0 q_1} \quad ; \quad (3.5.4)$$

Ví dụ có số liệu về hai loại hàng hoá tiêu thụ trên thị trường như sau:

Bảng 3.5.1: Giá và lượng hàng tiêu thụ tương ứng của hàng hoá

Loại hàng	Giá (Nghìn đồng)		Lượng hàng tiêu thụ (Kg)		Chỉ số giá đơn i_p	Chỉ số lượng hàng i_q
	Kỳ gốc	Kỳ n /cứu	Kỳ gốc	Kỳ n /cứu		
A	1	2	3	4	5=1:2	6=4:3
X	20	30	10	12	1,5	1,20
Y	4	8	30	20	2,0	0,67

Từ số liệu bảng 3.5.1

- Áp dụng công thức 3.5.3 có:

$$I_p = \frac{30 \times 10 + 8 \times 30}{20 \times 10 + 4 \times 30} = 1,688 \text{ hoặc } 168,8\%$$

- Áp dụng công thức 3.5.4 có:

$$I_p = \frac{30 \times 12 + 8 \times 20}{20 \times 12 + 4 \times 20} = 1,625 \text{ hoặc } 162,5\%$$

Các chỉ số theo Laspeyres và Paashe có logic tư duy khác nhau, đồng thời kết quả tính cũng có khác nhau. Thực ra không thể nói tính

theo công thức nào có ý nghĩa hơn công thức nào. Chỉ có điều quyền số của chỉ số theo Laspeyres là số liệu kỳ gốc nên thường thu thập thuận tiện hơn và sẽ đảm bảo kết quả tính toán kịp thời hơn. Mặt khác về trực quan người ta dễ nhận biết ý nghĩa của chỉ số này hơn, còn theo Paashe có ưu điểm là đảm bảo cơ cấu theo kỳ báo cáo nên sát với thực tế hơn.

- Nếu chọn quyền số kết hợp cả hai thời kỳ báo cáo và kỳ gốc, ta có chỉ số tổng hợp về giá cả theo Fisher:

$$I_p = \sqrt{\frac{\sum p_1 q_0}{\sum p_0 q_0} \times \frac{\sum p_1 q_1}{\sum p_0 q_1}} \quad ; \quad (3.5.5)$$

Chỉ số tổng hợp về giá cả theo Fisher là trung bình nhân của hai chỉ số tổng hợp về giá cả của Laspeyres và Paashe. Theo số liệu đã có, áp dụng công thức 3.5.5 có:

$$I_p = \sqrt{1,688 \times 1,625} = 1,656 \text{ hoặc } 165,6\%$$

Trong nhiều trường hợp tính toán với quyền số cố định ở các thời kỳ khác nhau theo phương pháp của Laspeyres và Paashe dẫn đến các kết quả quá sai lệch thì việc sử dụng chỉ số Fisher là cần thiết. Tuy nhiên, khả năng áp dụng và tính toán theo chỉ số của Fisher là khó khăn và phức tạp hơn.

** Chỉ số tổng hợp về giá cả theo không gian*

Trong phân tích so sánh kinh tế, có nhu cầu so sánh giá cả của một hoặc nhiều mặt hàng giữa các chợ trong một địa phương hoặc giữa các địa phương. Lúc này ta có các chỉ số giá cả theo không gian:

$$I_p = \frac{\sum p_A (q_A + q_B)}{\sum p_B (q_A + q_B)} = \frac{\sum p_A Q}{\sum p_B Q} \quad ; \quad (3.5.6)$$

Trong đó: A và B là hai địa phương cần so sánh ($q_A + q_B$) - quyền số của chỉ số. Đó là tổng khối lượng hàng tiêu thụ

của kỳ báo cáo và kỳ gốc của mỗi mặt hàng.

Ví dụ: Có tài liệu về giá cả và lượng hàng hoá tiêu thụ tại hai địa phương như sau:

Bảng 3.5.2: Giá và lượng hàng ở địa phương A và B

Mặt hàng	Địa phương A		Địa phương B	
	Giá cả (1000đ)	Lượng hàng bán ra (Kg)	Giá cả (1000đ)	Lượng hàng bán ra (Kg)
X	4,0	1000	3,5	1500
Y	2,0	2000	2,5	1000

Theo số liệu ở bảng 3.5.2, áp dụng công thức 3.5.6, ta tính được chỉ số giá cả địa phương A so với địa phương B như sau:

$$I_p = \frac{4 \times 2500 + 2 \times 3000}{3,5 \times 2500 + 2,5 \times 3000} = \frac{16000}{16250} = 0,9846 \text{ hoặc } 98,46\%$$

Như vậy, giá chung của cả hai mặt hàng ở địa phương A bằng 98,46% giá cả ở địa phương B, tức là giảm 1,54%.

b. Chỉ số tổng hợp về lượng hàng tiêu thụ

** Chỉ số tổng hợp về lượng hàng tiêu thụ theo thời gian*

- Nếu chọn quyền số là giá cả kỳ gốc, có chỉ số tổng hợp về lượng hàng tiêu thụ theo Laspeyres:

$$I_q = \frac{\sum p_0 q_1}{\sum p_0 q_0} \quad ; \quad (3.5.7)$$

- Nếu chọn quyền số là giá cả kỳ nghiên cứu, có chỉ số tổng hợp về lượng hàng tiêu thụ theo Paashe:

$$I_q = \frac{\sum p_1 q_1}{\sum p_1 q_0} \quad ; \quad (3.5.8)$$

Các chỉ số này cũng tiếp nối tư duy logic khác nhau của các chỉ số tổng hợp giá cả và kết quả tính toán theo hai công thức này cũng có sự khác nhau nhất định.

- Cũng như chỉ số tổng hợp về giá cả, Fisher đã đưa ra chỉ số tổng hợp về lượng hàng tiêu thụ với quyền số giá cả kết hợp của thời kỳ báo cáo và thời kỳ gốc:

Chỉ số tổng hợp về lượng hàng của Fisher cũng là trung bình nhân của hai chỉ số tổng hợp về lượng hàng tiêu thụ theo Laspeyres và Paashe:

$$I_q = \sqrt{\frac{\sum p_0 q_1}{\sum p_0 q_0} \times \frac{\sum p_1 q_1}{\sum p_1 q_0}} \quad ; \quad (3.5.9)$$

Theo số liệu đã cho ở bảng 3.5.1 tính được:

- Theo công thức 3.5.7:

$$I_q = \frac{20 \times 12 + 4 \times 20}{20 \times 10 + 4 \times 30} = \frac{320}{320} = 1,00 \text{ hoặc } 100,0\%$$

- Theo công thức 3.5.8:

$$I_q = \frac{30 \times 12 + 8 \times 20}{30 \times 10 + 4 \times 20} = \frac{520}{540} = 0,963 \text{ hoặc } 96,3\%$$

- Theo công thức 3.5.9:

$$I_q = \sqrt{1,00 \times 0,963} = 0,981 \text{ hoặc } 98,1\%$$

** Chỉ số tổng hợp về lượng hàng theo không gian*

Chỉ số tổng hợp lượng hàng tiêu thụ theo không gian có thể dùng giá so sánh tính thống nhất cho các địa bàn:

$$I_q = \frac{\sum p_s q_A}{\sum p_s q_B} \quad ; \quad (3.5.10)$$

Trong đó:

A và B là hai địa phương cần so sánh,

p_s là giá so sánh của từng mặt hàng.

Tuy nhiên, trong nhiều trường hợp, ta không có giá so sánh cho tất cả các mặt hàng, nên cần sử dụng giá bình quân của hai địa phương cần so sánh:

$$\bar{p} = \frac{p_A q_A + p_B q_B}{q_A + q_B}$$

Và chỉ số tổng hợp lúc này là:

$$I_q = \frac{\Sigma \bar{p} \cdot q_A}{\Sigma \bar{p} \cdot q_B} \quad ; \quad (3.5.11)$$

Theo số liệu ở bảng 3.5.2, ta có:

- Giá bình quân 1kg hàng X:

$$\bar{p}_X = \frac{4 \times 1000 + 3,5 \times 1500}{1000 + 1500} = \frac{9250}{2500} = 3,7 \text{ (nghìn đồng)}$$

- Giá bình quân 1kg hàng Y:

$$\bar{p}_Y = \frac{2 \times 2000 + 2,5 \times 1000}{2000 + 1000} = \frac{6500}{3000} = 2,166 \text{ (nghìn đồng)}$$

Áp dụng công thức 3.5.11 ta tính được chỉ số lượng hàng tiêu thụ giữa địa phương A so với địa phương B:

$$I_q = \frac{3,7 \times 1000 + 2,166 \times 2000}{3,7 \times 1500 + 2,166 \times 1000} = \frac{8032}{7716} = 1,041 \text{ hoặc } 104,1\%$$

Như vậy, lượng hàng hoá địa phương A bằng 104,1% lượng hàng hoá địa phương B, tức là cao hơn 4,1%.

3.5.3. Chỉ số bình quân

Chỉ số bình quân là một dạng biến đổi của chỉ số tổng hợp, công thức tính được trình bày dưới dạng một số bình quân. Có hai loại chỉ số bình quân:

a. Chỉ số bình quân số học gia quyền - dạng biến đổi từ một số chỉ số tổng hợp có quyền số cố định ở thời kỳ gốc.

- Chỉ số tổng hợp về giá cả có quyền số là lượng hàng tiêu thụ cố định ở thời kỳ gốc (công thức 3.5.3 - Laspeyres):

$$I_p = \frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} = \frac{\Sigma \frac{p_1}{p_0} p_0 q_0}{\Sigma p_0 q_0} = \frac{\Sigma i_p \cdot p_0 q_0}{\Sigma p_0 q_0} \quad ; \quad (3.5.12)$$

- Chỉ số tổng hợp về khối lượng hàng hoá có quyền số cố định ở thời kỳ gốc (công thức 3.5.7 - Laspeyres):

$$I_q = \frac{\Sigma p_0 q_1}{\Sigma p_0 q_0} = \frac{\Sigma \frac{q_1}{q_0} p_0 q_0}{\Sigma p_0 q_0} = \frac{\Sigma i_q \cdot p_0 q_0}{\Sigma p_0 q_0} \quad ; \quad (3.5.13)$$

Trong đó: $i_p = \frac{p_1}{p_0}$ và $i_q = \frac{q_1}{q_0}$ là các chỉ số cá thể về giá và lượng hàng hoá tiêu thụ. Ở đây, các chỉ số cá thể đóng vai trò là lượng biến và $p_0 q_0$ là quyền số của chỉ số tổng hợp được cố định ở thời kỳ gốc.

Từ số liệu bảng 3.5.1:

- Áp dụng công thức 3.5.12 ta có chỉ số giá:

$$I_p = \frac{1,5 \times 200 + 2 \times 120}{200 + 120} = 1,688 \text{ hoặc } 168,8\%$$

- Áp dụng công thức 3.5.13 ta có chỉ số lượng hàng hoá tiêu thụ:

$$I_q = \frac{1,2 \times 200 + 0,67 \times 120}{200 + 120} = 1,000 \text{ hoặc } 100,0\%$$

b. Chỉ số bình quân điều hoà gia quyền - dạng biến đổi từ một số chỉ số tổng hợp có quyền số cố định ở thời kỳ báo cáo.

- Chỉ số tổng hợp về giá có quyền số là lượng hàng hoá tiêu thụ cố định ở thời kỳ báo cáo (công thức 3.5.4 - Paasche):

$$I_p = \frac{\sum p_1 q_1}{\sum \frac{p_0}{p_1} p_1 q_1} = \frac{\sum p_1 q_1}{\sum \frac{1}{i_p} p_1 q_1} ; \quad (3.5.14)$$

- Chỉ số tổng hợp về lượng hàng hoá tiêu thụ có quyền số là giá cả cố định ở thời kỳ báo cáo (công thức 3.5.8 - Paasche):

$$I_q = \frac{\sum p_1 q_1}{\sum \frac{q_0}{q_1} p_1 q_1} = \frac{\sum p_1 q_1}{\sum \frac{1}{i_q} p_1 q_1} ; \quad (3.5.15)$$

Trong đó các chỉ số cá thể i_p và i_q đóng vai trò lượng biến và $p_1 q_1$ là quyền số của chỉ số bình quân chung.

Cũng từ số liệu bảng 3.5.1:

- Áp dụng công thức 3.5.14 ta có chỉ số giá:

$$I_p = \frac{320 + 160}{\frac{320}{1,5} + \frac{160}{2,0}} = 1,636 \text{ hoặc } 163,6\%$$

- Áp dụng công thức 3.5.15 ta có chỉ số lượng hàng hoá:

$$I_q = \frac{320 + 160}{\frac{320}{1,2} + \frac{160}{2,0}} = 0,963 \text{ hoặc } 96,3\%$$

Các chỉ số bình quân được áp dụng trong các trường hợp có tài liệu về các chỉ số cá thể và đặc biệt có ý nghĩa khi tiếp tục biến đổi quyền số của chỉ số về dạng "tỷ trọng giá trị của từng loại hàng hoá"

để có thể sử dụng thuận lợi tỷ trọng đó khi tính toán và trong những trường hợp cần thiết có thể dùng tỷ trọng tương ứng để thay thế.

3.5.4. Chỉ số liên hoàn và chỉ số định gốc

3.5.4.1. Chỉ số liên hoàn

Chỉ số liên hoàn là chỉ số tính cho nhiều thời kỳ liên tiếp nhau, trong đó mỗi chỉ số đều so sánh thời kỳ nghiên cứu với thời kỳ liền kề trước đó. Thời kỳ quyền số của các chỉ số liên hoàn có thể thay đổi (trường hợp này gọi là quyền số khả biến) hoặc không thể thay đổi (trường hợp này gọi là quyền số bất biến).

- *Chỉ số liên hoàn với quyền số khả biến*: Ví dụ, chỉ số giá bán lẻ các mặt hàng tính cho tháng 2, 3, 4 (chỉ số giá tháng 2 so với tháng 1 lấy quyền số là lượng hàng tháng 2, chỉ số giá tháng 3 so với tháng 2 lấy quyền số là lượng hàng tháng 3 và chỉ số giá tháng 4 so với tháng 3 lấy quyền số là lượng hàng tháng 4).

$$I_{p2/1} = \frac{\sum p_2 q_2}{\sum p_1 q_2} ; I_{p3/2} = \frac{\sum p_3 q_3}{\sum p_2 q_3} ; I_{p4/3} = \frac{\sum p_4 q_4}{\sum p_3 q_4} ; \quad (3.5.16)$$

- *Chỉ số liên hoàn với quyền số bất biến*: Ví dụ, chỉ số khối lượng sản phẩm công nghiệp tính cho tháng 2, 3, 4 với cùng giá so sánh hoặc giá cố định của sản phẩm (giá năm 1994) ký hiệu là p_s .

$$I_{q2/1} = \frac{\sum p_s q_2}{\sum p_s q_1} ; I_{q3/2} = \frac{\sum p_s q_3}{\sum p_s q_2} ; I_{q4/3} = \frac{\sum p_s q_4}{\sum p_s q_3} ; \quad (3.5.17)$$

Quyền số bất biến của chỉ số tuy có cơ cấu khác nhiều hơn so với thực tế, nhưng có tính khả thi cao hơn vì nhiều năm mới phải xác định giá một lần. Trong nhiều trường hợp thực tế đã không thể áp dụng được quyền số khả biến, mà phải thay bằng quyền số bất biến. Ví dụ: Chỉ số khối lượng sản phẩm công nghiệp dùng quyền số là giá cố định (giá của một năm nào đó được chọn để tính toán thống nhất cho nhiều năm); chỉ số giá tiêu dùng dùng quyền số là tỷ trọng khối lượng hàng

hoá tiêu dùng (tỷ trọng hàng hoá của một năm nào đó chọn để tính toán thống nhất cho một số năm).

3.5.4.2. Chỉ số định gốc

Chỉ số định gốc là chỉ số tính cho nhiều thời kỳ khác nhau so với một thời kỳ được chọn làm gốc cố định. Thời kỳ quyền số của các chỉ số định gốc có thể thay đổi (trường hợp này gọi là quyền số khả biến) hoặc không thay đổi (trường hợp này gọi là quyền số bất biến).

- *Chỉ số định gốc với quyền số khả biến:* Ví dụ, chỉ số giá bán lẻ các tháng 2, 3, 4 so với tháng 1.

$$I_{p2/1} = \frac{\sum p_2 q_2}{\sum p_1 q_2}; I_{p3/2} = \frac{\sum p_3 q_3}{\sum p_1 q_3}; I_{p4/3} = \frac{\sum p_4 q_4}{\sum p_1 q_4}; \quad (3.5.18)$$

- *Chỉ số định gốc với quyền số bất biến:* Ví dụ, chỉ số khối lượng sản phẩm công nghiệp các tháng 2, 3, 4 so với tháng 1, tính theo giá so sánh hoặc giá cố định của sản phẩm (giá năm 1994):

$$I_{q2/1} = \frac{\sum p_s q_2}{\sum p_s q_1}; I_{q3/2} = \frac{\sum p_s q_3}{\sum p_s q_1}; I_{q4/3} = \frac{\sum p_s q_4}{\sum p_s q_1}; \quad (3.5.19)$$

Giữa chỉ số định gốc và chỉ số liên hoàn (với quyền số bất biến) có quan hệ sau: Tích các chỉ số liên hoàn bằng chỉ số định gốc trong thời kỳ đó. Ví dụ: Chỉ số liên hoàn và chỉ số định gốc về khối lượng sản phẩm công nghiệp:

$$\frac{\sum p_n q_2}{\sum p_n q_1} \times \frac{\sum p_n q_3}{\sum p_n q_2} \times \frac{\sum p_n q_4}{\sum p_n q_3} = \frac{\sum p_n q_4}{\sum p_n q_1}$$

$$\text{hoặc} \quad I_{q2/1} \times I_{q3/2} \times I_{q4/3} = I_{q4/1}; \quad (3.5.20)$$

3.5.5. Chỉ số sản phẩm so sánh được và sản phẩm không so sánh được

Trong thực tế sản xuất và tiêu thụ sản phẩm, hàng hóa (từ đây gọi chung là sản phẩm), ngoài những loại cùng sản xuất và tiêu thụ ở cả hai thời kỳ (kỳ gốc và kỳ báo cáo) gọi là "sản phẩm so sánh được", còn có những loại sản phẩm chỉ sản xuất hoặc tiêu thụ ở một trong hai thời kỳ đó gọi là "sản phẩm không so sánh được" (xem ví dụ bảng 3.5.3).

Bảng 3.5.3: Số liệu và đơn giá thực tế một số loại sản phẩm sản xuất trong năm 2003 và 2004 ⁽¹⁾ của công ty "A"

Tên sản phẩm	Đơn vị tính SP	Khối lượng sản phẩm		Đơn giá (1000đ)		Giá trị sản xuất (Triệu đồng)	
		Kỳ gốc (q ₀)	Kỳ báo cáo (q ₁)	Kỳ gốc (p ₀)	Kỳ báo cáo (p ₁)	Kỳ gốc (p ₀ q ₀)	Kỳ báo cáo (p ₁ q ₁)
Sản phẩm 1	1000V	10.000	12.000	238	240	2.380	2.880
Sản phẩm 2	1000C	20.000	21.000	550	500	11.000	10.500
Sản phẩm 3	Mét	5.000	7.000	35	38	175	266
Sản phẩm 4	Tấm	-	3.800	-	1.000	-	3.800
Sản phẩm 5	Tấm	2.200	-	1.200	-	2.640	-
Tổng cộng	x	x	x	x	x	16.195	17.446

Số liệu bảng 3.5.3 cho thấy công ty "A" sản xuất 5 loại sản phẩm, có 3 loại sản phẩm 1, 2, 3 được sản xuất ở cả hai năm (2003 và 2004) và đó là những sản phẩm so sánh được, còn sản phẩm thứ 4 chỉ sản xuất ở năm 2004 (năm báo cáo) và sản phẩm thứ 5 chỉ sản xuất ở năm 2003 (năm gốc) là những sản phẩm không so sánh được. Trường hợp như trên thì sẽ tính chỉ số khối sản phẩm như thế nào?

Như ta đã biết chỉ số khối lượng sản phẩm không chỉ phản ánh sự

⁽¹⁾ Năm 2003 là năm gốc và năm 2004 là năm báo cáo.

tăng lên của những sản phẩm đã có ở thời kỳ trước, mà còn phải phản ánh cả sự thay đổi về mặt hàng sản xuất ra (sự tăng thêm hay giảm bớt mặt hàng sản xuất cũng chính là sự tăng lên hay giảm đi của khối lượng sản phẩm sản xuất ra).

Nếu áp dụng đơn thuần công thức tính chỉ số khối lượng sản phẩm với quyền số là giá cả thời kỳ gốc (theo Laspayres) hoặc với quyền số là giá cả thời kỳ báo cáo (theo Paasche) đều chỉ tính được cho các sản phẩm so sánh được (ở trên sản phẩm 1, 2 và 3), còn các loại sản phẩm không so sánh được (4 và 5) đều không đủ thông tin để tính toán (hoặc là thiếu số liệu kỳ gốc, hoặc là thiếu số liệu kỳ báo cáo).

Vấn đề là phải xây dựng được chỉ số để áp dụng cho cả trường hợp có sản phẩm không so sánh được.

* *
*

Khi sản xuất có cả sản phẩm so sánh được và sản phẩm không so sánh được thì giá trị sản xuất theo giá thực tế được viết dưới dạng:

$$\sum pq = (\sum p'q' + \sum p''q'') \quad ; \quad (3.5.21a)$$

Trong đó:

$\sum pq$ - Giá trị sản xuất của toàn bộ sản phẩm sản xuất với p là giá cả và q là khối lượng từng loại sản phẩm;

$\sum p'q'$ - Giá trị sản xuất của những loại sản phẩm so sánh được với p' là giá cả và q' là khối lượng sản phẩm tương ứng;

$\sum p''q''$ - Giá trị sản xuất của những loại sản phẩm không so sánh được với p'' là giá cả và q'' là khối lượng sản phẩm tương ứng.

Tiếp tục biến đổi công thức 3.5.21a:

$$\sum pq = (\sum p'q' + \sum p''q'') = \sum p'q' \times \frac{(\sum p'q' + \sum p''q'')}{\sum p'q'}$$

$$= \sum p'q' \times \frac{\sum pq}{\sum p'q'} = (\sum p'q') \times K \quad ; \quad (3.5.21b)$$

Trong đó: K là tỷ số giữa giá trị sản xuất theo giá thực tế của toàn bộ sản phẩm (kể cả sản phẩm so sánh được và sản phẩm không so sánh được) và giá trị sản xuất của những sản phẩm so sánh được. Ở đây K tạm gọi là "*Hệ số thay đổi mặt hàng sản xuất*".

Trên cơ sở công thức 3.5.21b có thể xây dựng được các chỉ số sau:

a. Chỉ số khối lượng sản phẩm so sánh được (I'_q)

$$I'_q = \frac{\sum p'_0 q'_1}{\sum p'_0 q'_0} \quad ; \quad (3.5.22)$$

Chỉ số khối lượng sản phẩm so sánh được theo công thức 3.5.22 (viết gọn là chỉ số sản phẩm so sánh được) là dạng cơ bản của chỉ số khối lượng theo Laspeyres.

Từ số liệu bảng 3.5.3 áp dụng công thức 3.5.22 tính được:

$$I'_q = \frac{(238 \times 12000) + (550 \times 21000) + (35 \times 7000)}{(238 \times 10000) + (550 \times 20000) + (35 \times 5000)} = 1,0808$$

hoặc 108,08%

b. Chỉ số khối lượng sản phẩm không so sánh được (I''_q)

$$I''_q = \frac{K_1}{K_0} \quad ; \quad (3.5.23)$$

Trong đó: K_1, K_0 - Hệ số thay đổi mặt hàng sản xuất.

I''_q - Chỉ số khối lượng sản phẩm không so sánh được (viết gọn là chỉ số sản phẩm không so sánh được) phản ánh biến động khối lượng sản phẩm do mở rộng hay thu hẹp mặt hàng sản xuất. Nếu $I_k > 1$ nghĩa là kỳ báo cáo có khối lượng mặt hàng mới xuất hiện lớn hơn khối lượng mặt hàng cũ mất đi và được gọi là trường hợp mở rộng mặt

hàng sản xuất; nếu $I_k < 1$ nghĩa là kỳ báo cáo có khối lượng mặt hàng mới xuất hiện nhỏ hơn khối lượng mặt hàng cũ mất đi và được gọi là trường hợp thu hẹp mặt hàng sản xuất. Còn nếu $I_k = 1$ thì hoặc là không có mặt hàng mới xuất hiện và cũng không có mặt hàng cũ mất đi, hoặc là có cả mặt hàng mới xuất hiện và mặt hàng cũ mất đi nhưng tỷ trọng giá trị của những mặt hàng không so sánh được chiếm trong tổng giá trị sản xuất ở thời kỳ báo cáo và thời kỳ gốc tương đương như nhau.

Từ số liệu bảng 3.5.3 ta tính được:

+ *Hệ số K*:

- Năm 2003

$$K_q = \frac{16195}{2380 + 11000 + 175} = \frac{16195}{13555} = 1,195$$

- Năm 2004

$$K_q = \frac{17446}{2880 + 10500 + 266} = \frac{17446}{13646} = 1,2785$$

+ *Chỉ số sản phẩm không so sánh được (áp dụng công thức 3.5.23)*:

$$I''_q = \frac{1,2785}{1,1932} = 1,0715 \text{ hoặc } 107,15\%$$

c. *Chỉ số khối lượng sản phẩm*

Nhân 2 chỉ số sản phẩm so sánh được (I'_q) và chỉ số sản phẩm không so sánh được (I''_q) ta được chỉ số khối lượng sản phẩm (I_q):

$$I'_q \times I''_q = I_q \quad ; \quad (3.5.24)$$

Theo số liệu tính được ở mục a và b, áp dụng công thức 3.5.24 ta có:

$$1,0808 \times 1,0715 = 1,1581 \text{ hoặc } 115,81\%$$

Như vậy khối lượng sản phẩm sản xuất của công ty "A" năm 2004 so với năm 2003 tăng 15,81%; trong đó do sản phẩm so sánh

được tăng làm tăng 8,08% và mở rộng mặt hàng sản xuất làm tăng 7,15%.

3.5.6. *Hệ thống chỉ số*

Hệ thống chỉ số là dãy các chỉ số có liên hệ với nhau, hợp thành một đẳng thức nhất định. Có nhiều loại hệ thống chỉ số, trong thực tế công tác thống kê thường gặp hai loại: hệ thống chỉ số tổng hợp và hệ thống chỉ số nghiên cứu biến động chỉ tiêu bình quân.

3.5.6.1. *Hệ thống chỉ số tổng hợp*

Trở lại số liệu ở bảng 3.5.1, nếu lấy tổng giá trị hàng hoá tiêu thụ kỳ báo cáo ($\Sigma p_1 q_1$) chia cho tổng giá trị hàng hoá ở kỳ gốc ($\Sigma p_0 q_0$) ta được chỉ số giá trị (I_{pq}). Nghiên cứu mối quan hệ giữa chỉ số giá trị với các chỉ số giá cả (I_p) và chỉ số lượng hàng hoá tiêu thụ (I_q), ta có:

$$\begin{aligned} \text{Chỉ số giá trị} &= \text{Chỉ số giá} \times \text{Chỉ số lượng} \\ I_{pq} &= I_p \times I_q \quad ; \quad (3.5.25) \end{aligned}$$

Tuy nhiên, do các cách xây dựng chỉ số giá cả và chỉ số lượng hàng theo những quy định khác nhau, nên ta cũng có các hệ thống chỉ số khác nhau.

a. *Nếu chỉ số giá theo Paashe và chỉ số khối lượng theo Laspayres thì ta có hệ thống chỉ số:*

$$\frac{\Sigma p_1 q_1}{\Sigma p_0 q_0} = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1} \times \frac{\Sigma p_0 q_1}{\Sigma p_0 q_0} \quad ; \quad (3.5.26)$$

b. *Nếu chỉ số giá theo Laspayres và chỉ số khối lượng theo Paashe thì ta có hệ thống chỉ số:*

$$\frac{\Sigma p_1 q_1}{\Sigma p_0 q_0} = \frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \frac{\Sigma p_1 q_1}{\Sigma p_1 q_0} \quad ; \quad (3.5.27)$$

Hai hệ thống trên không cho ta đẳng thức để đảm bảo quan hệ tích số đã nêu theo đẳng thức 3.5.25:

Theo công thức của Fisher, ta có đẳng thức:

$$\frac{\sum p_1 q_1}{\sum p_0 q_0} = \sqrt{\frac{\sum p_1 q_0}{\sum p_0 q_0} \times \frac{\sum p_1 q_1}{\sum p_0 q_1}} \times \sqrt{\frac{\sum p_1 q_1}{\sum p_1 q_0} \times \frac{\sum p_0 q_1}{\sum p_0 q_0}} ; \quad (3.5.28)$$

Công thức này đảm bảo quan hệ tích số như đẳng thức 3.5.25, nhưng điều kiện áp dụng và tính toán khá phức tạp, vì phải hai lần tính lại theo quyền số.

Về mặt lý thuyết thống kê xã hội chủ nghĩa nói chung cũng như thống kê nước ta nói riêng đã sử dụng hệ thống chỉ số (3.5.25) tức là trong hệ thống chỉ số có chỉ số giá tổng hợp là theo Paashe, còn chỉ số tổng hợp khối lượng hàng hoá tiêu thụ là theo Laspeyres. Tuy nhiên, trong thực tế công tác thống kê, tùy theo điều kiện cụ thể của mỗi loại chỉ số khác nhau mà có những quy định thời kỳ lựa chọn quyền số cho thích hợp.

Hệ thống chỉ số tổng hợp được dùng để phân tích ảnh hưởng của các nhân tố cấu thành đối với một hiện tượng phức tạp, cho ta các thông tin mới về sự biến động của hiện tượng theo sự tác động của các nhân tố cấu thành đó. Vì vậy, hệ thống này còn được dùng cho nhiều quan hệ khác, chẳng hạn:

$$\begin{array}{ccccc} \text{Số sản phẩm} & & \text{Năng suất lao động} & & \text{Số} \\ \text{sản xuất} & = & \text{của 1 công nhân} & \times & \text{công nhân} \end{array}$$

$$\begin{array}{ccccc} \text{Giá thành toàn} & & \text{Giá thành bình quân} & & \text{Số sản phẩm} \\ \text{bộ sản phẩm} & = & \text{một sản phẩm} & \times & \text{sản xuất} \end{array}$$

v.v.,...

Hệ thống này cũng được sử dụng trong phân tích mức độ hoàn thành kế hoạch của một doanh nghiệp, của một vùng lãnh thổ (tỉnh, huyện,...).

$$\begin{array}{ccccc} \text{Chỉ số} & & \text{Chỉ số nhiệm vụ} & & \text{Chỉ số hoàn thành} \\ \text{phát triển} & = & \text{kế hoạch} & \times & \text{kế hoạch} \end{array}$$

Tức là:

$$\frac{\sum p_1 q_1}{\sum p_0 q_0} = \frac{\sum p_k q_k}{\sum p_0 q_0} \times \frac{\sum p_1 q_1}{\sum p_k q_k} ; \quad (3.5.29)$$

Với k - Thời kỳ kế hoạch.

3.5.6.2. Hệ thống chỉ số nghiên cứu biến động chỉ tiêu bình quân

Khi nghiên cứu biến động chỉ tiêu bình quân có 3 chỉ số lập thành một hệ thống: Chỉ số cấu thành khả biến, chỉ số cấu thành cố định và chỉ số ảnh hưởng kết cấu.

a. *Chỉ số cấu thành khả biến.* Đó là chỉ tiêu tương đối biểu hiện quan hệ so sánh giữa hai mức độ bình quân của hiện tượng nghiên cứu. Muốn tính chỉ số này, trước hết cần tính mức độ bình quân của hiện tượng ở hai thời kỳ, rồi đem so sánh hai mức độ đó với nhau. Công thức tính:

$$I_{\bar{x}} = \frac{\bar{x}_1}{\bar{x}_0} = \frac{\sum x_1 f_1}{\sum f_1} : \frac{\sum x_0 f_0}{\sum f_0} ; \quad (3.5.30)$$

Trong đó:

$I_{\bar{x}}$ - Chỉ số cấu thành khả biến;

$\bar{x}_1; \bar{x}_0$ - Mức độ bình quân kỳ báo cáo và kỳ gốc;

f_1, f_0 - Quyền số của số bình quân kỳ báo cáo và kỳ gốc.

Chỉ số cấu thành khả biến phản ánh sự biến động đồng thời của hai nhân tố: Tiêu thức bình quân hoá và kết cấu tổng thể. Do đó, chỉ số cấu thành khả biến có thể được phân tích thành hai chỉ số nhân tố: Chỉ số cấu thành cố định và chỉ số ảnh hưởng kết cấu.

Trong phân tích thống kê chỉ số cấu thành khả biến thường được dùng để biểu hiện sự biến động một cách tổng quát của các chỉ tiêu bình quân như: Biến động giá thành bình quân, biến động năng suất lao động bình quân, biến động năng suất thu hoạch bình quân, v.v...

b. *Chỉ số cấu thành cố định.* Đó là chỉ tiêu tương đối nêu lên ảnh hưởng biến động của riêng tiêu thức bình quân hoá đối với sự biến động của chỉ tiêu bình quân. Trong chỉ số này kết cấu của tổng thể được cố định ở một kỳ nhất định.

Nếu chỉ số cấu thành cố định tính theo kết cấu tổng thể kỳ báo cáo:

$$I_{\bar{x}} = \frac{\sum x_1 f_1}{\sum f_1} : \frac{\sum x_0 f_1}{\sum f_1} \quad ; \quad (3.5.31a)$$

sau khi giản ước ta có:

$$I_{\bar{x}} = \frac{\sum x_1 f_1}{\sum x_0 f_1} \quad ; \quad (3.5.31b)$$

Trong đó:

$I_{\bar{x}}$ - Chỉ số cấu thành cố định;

$x_1; x_0$ - Lượng biến kỳ báo cáo và kỳ gốc của chỉ tiêu bình quân;

$\frac{f_1}{\sum f_1}$ - Kết cấu của tổng thể kỳ báo cáo.

Chỉ số cấu thành cố định được dùng để phân tích chất lượng của các công tác sản xuất, quản lý kinh tế, như: Đánh giá ảnh hưởng biến động của bản thân yếu tố giá thành sản phẩm đối với biến động của giá thành bình quân, đánh giá ảnh hưởng biến động của bản thân yếu tố tiền lương đối với biến động của tiền lương bình quân,...

c. *Chỉ số ảnh hưởng kết cấu.* Đó là chỉ tiêu tương đối phân tích ảnh hưởng biến động của kết cấu tổng thể đối với sự biến động của chỉ tiêu bình quân. Trong chỉ số này, tiêu thức bình quân hoá được cố định ở một kỳ nhất định.

Nếu cố định tiêu thức bình quân hoá ở kỳ gốc thì chỉ số ảnh hưởng kết cấu có dạng:

$$I_{f/\Sigma f} = \frac{\sum x_0 f_1}{\sum f_1} : \frac{\sum x_0 f_0}{\sum f_0} \quad ; \quad (3.5.32)$$

Trong đó:

$I_{f/\Sigma f}$ - Chỉ số cấu thành kết cấu;

x_0 - Lượng biến kỳ gốc của chỉ tiêu bình quân;

$\frac{f_1}{\sum f_1}; \frac{f_0}{\sum f_0}$ - Kết cấu của tổng thể kỳ báo cáo và kỳ gốc.

Chỉ số ảnh hưởng kết cấu thường được dùng để phân tích ảnh hưởng của nhân tố kết cấu đối với biến động của các chỉ tiêu bình quân như: Thay đổi kết cấu sản phẩm cùng loại nhưng có giá thành khác nhau đối với sự thay đổi của giá thành bình quân, thay đổi kết cấu công nhân có mức lương khác nhau đối với sự thay đổi tiền lương bình quân,...

3.6. PHƯƠNG PHÁP CÂN ĐỐI

Phương pháp cân đối là một phương pháp chính lý và phân tích các số liệu thống kê bằng cách sử dụng các bảng cân đối để nghiên cứu các quan hệ tỷ lệ, các mối liên hệ qua lại giữa các hiện tượng và quá trình kinh tế - xã hội và để so sánh đối chiếu số liệu thu được từ nhiều nguồn và phân tổ theo nhiều tiêu thức khác nhau.

Bảng cân đối là một hình thức trình bày kết cấu của cùng một tổng thể (hiện tượng hoặc quá trình kinh tế - xã hội theo hai giác độ khác nhau) để phản ánh các quan hệ cân đối giữa các bộ phận trong tổng thể hoặc để so sánh, kiểm tra số liệu đã thu thập được từ nhiều nguồn khác nhau.

Trong thống kê, các chỉ tiêu của bảng cân đối có thể biểu hiện bằng đơn vị hiện vật hoặc đơn vị giá trị, đơn vị thời gian lao động. Do đó, phương pháp cân đối được sử dụng rộng rãi trong nhiều bộ môn

thống kê kinh tế để phản ánh và kiểm tra quan hệ cân đối giữa sản xuất và tiêu dùng, giữa thu và chi ngân sách, giữa tích lũy và tiêu dùng, giữa sản xuất và tiêu thụ sản phẩm, v.v... Dựa vào sự cân bằng của phương trình kinh tế trong bảng cân đối, có thể phát hiện các mặt mất cân đối, các sai sót trong số liệu thống kê.

Trong thống kê thường sử dụng hai loại bảng cân đối.

3.6.1. Bảng cân đối "đơn"

Đó là loại bảng cân đối biểu hiện một tổng thể gồm hai phần tử tương ứng với hai mặt đối lập, trong đó mỗi phần được phân tổ theo các tiêu thức khác nhau. Các loại bảng cân đối đơn thường gặp như cân đối xuất nhập khẩu hàng hoá, cân đối giữa sản xuất và tiêu dùng, cân đối giữa nguồn và sử dụng lao động, v.v... Cấu trúc của bảng cân đối đơn được trình bày theo dòng hoặc theo cột. Ví dụ, bảng cân đối lao động xã hội có dạng sau:

Bảng 3.6.1. Bảng cân đối lao động xã hội

Phần A. Nguồn lao động	Ký hiệu	Phần B. Sử dụng lao động	Ký hiệu
1. Lao động trong độ tuổi lao động	A ₁	1. Lao động làm việc trong các ngành kinh tế	B ₁
2. Lao động ngoài độ tuổi lao động	A ₂	2. Lao động dự trữ	B ₂

Cộng	$\sum_{i=1}^n A_i$	Cộng	$\sum_{j=1}^m B_j$
------	--------------------	------	--------------------

Phương trình kinh tế của loại bảng cân đối này có dạng:

$$\sum_{i=1}^n A_i = \sum_{j=1}^m B_j \qquad ; \qquad (3.6.1)$$

Trong đó:

A_i và $\sum_{i=1}^n A_i$ - Bộ phận thứ i và tổng n các bộ phận của phần thứ nhất (i chỉ thứ tự các bộ phận với i = 1,2,...n);

B_j và $\sum_{j=1}^m B_j$ - Bộ phận thứ j và tổng m các bộ phận của phần thứ hai (j chỉ thứ tự các bộ phận với j = 1,2,...m).

3.6.2. Bảng cân đối "kép"

Bảng cân đối "kép" (còn gọi là cân đối "bàn cờ") là loại bảng cân đối biểu hiện một tổng thể gồm hai phần tử tương ứng với hai mặt đối lập, trong đó mỗi bộ phận trong kết cấu của phần thứ nhất được phân tổ theo kết cấu của phần thứ hai và ngược lại mỗi bộ phận trong kết cấu của bộ phận thứ hai cũng được phân tổ theo kết cấu của phần thứ nhất.

Về cấu trúc, bảng cân đối kép được trình bày dưới dạng cân đối bàn cờ kết hợp giữa dòng và cột. Mỗi cột đều chia theo tất cả các dòng và mỗi dòng cũng được chia theo tất cả các cột.

Ví dụ: Bảng cân đối nguồn vốn và sử dụng cho hoạt động y tế quốc gia. Bảng cân đối này có hai phần: Nguồn vốn - trình bày theo cột và sử dụng vốn theo các loại hình hoạt động y tế - trình bày theo dòng, được phân tổ như sau:

Bảng 3.6.2: Bảng cân đối nguồn vốn và sử dụng cho hoạt động y tế quốc gia

Sử dụng vốn \ Nguồn vốn	Ngân sách nhà nước	BHXH và BHYT	...	Nguồn vốn khác	Tổng nguồn vốn
Phòng bệnh, phòng dịch	a_{11}	a_{12}	...	a_{1m}	$\sum_{j=1}^m a_{1j}$
Khám chữa bệnh	a_{21}	a_{22}	...	a_{2m}	$\sum_{j=1}^m a_{2j}$
.....		...	...	...	...
Hoạt động y tế khác	a_{n1}	a_{n2}	...	a_{nm}	$\sum_{j=1}^m a_{nj}$
Tổng sử dụng vốn	$\sum_{i=1}^n a_{i1}$	$\sum_{i=1}^n a_{i2}$	...	$\sum_{i=1}^n a_{im}$	$\sum_{j=1}^m \sum_{i=1}^n a_{ij}$

Phương trình kinh tế của bảng cân đối kép có dạngP:

$$\sum_{i=1}^n \sum_{j=1}^m a_{ij} = \sum_{j=1}^m \sum_{i=1}^n a_{ij} \quad ; \quad (3.6.2)$$

Trong đó:

$\sum_{j=1}^m a_{ij}$ - Từng hoạt động i theo tổng các nguồn của j;

$\sum_{i=1}^n a_{ij}$ - Từng nguồn j theo tất cả các hoạt động i.

PHẦN BỐN

MỘT SỐ CHỈ TIÊU CHỦ YẾU TRONG THỐNG KÊ TÀI KHOẢN QUỐC GIA

Hệ thống tài khoản quốc gia là một tập hợp đầy đủ, phù hợp và linh hoạt các tài khoản kinh tế vĩ mô, xây dựng trên những khái niệm, định nghĩa, quy tắc hạch toán được thừa nhận trên phạm vi quốc tế ⁽¹⁾. Mục đích của việc thiết kế tài khoản quốc gia nhằm phục vụ cho nhu cầu quản lý, phân tích và hoạch định chính sách kinh tế. Những chỉ tiêu kinh tế tổng hợp trong tài khoản quốc gia có mối liên hệ mật thiết với nhau, phản ánh "kết quả" hoạt động của nền kinh tế từ sản xuất, thu nhập, phân phối lại thu nhập, đến tiêu dùng, để dành, tích lũy tài sản và của cải của nền kinh tế. Phần này sẽ đề cập tới nội dung và ý nghĩa kinh tế của một số chỉ tiêu kinh tế tổng hợp trong Hệ thống tài khoản quốc gia. Trước khi đề cập từng chỉ tiêu, chúng ta điểm lại một số khái niệm cơ bản trong thống kê tài khoản quốc gia.

4.1. MỘT SỐ KHÁI NIỆM CƠ BẢN

4.1.1. Sản xuất

Trong tài khoản quốc gia sản xuất được định nghĩa như sau: "Sản xuất là quá trình sử dụng lao động và máy móc thiết bị của các đơn vị thể chế để chuyển những chi phí là vật chất và dịch vụ thành sản phẩm là vật chất và dịch vụ khác. Tất cả những hàng hóa và dịch vụ được sản xuất ra phải có khả năng bán trên thị trường hay ít ra cũng có khả năng cung cấp cho một đơn vị thể chế khác có thu tiền hoặc không thu tiền" ⁽¹⁾. Với khái niệm sản xuất, cần lưu ý một số đặc trưng sau:

⁽¹⁾ Mục 1.1, Tài khoản quốc gia 1993 của Liên Hợp Quốc.

⁽¹⁾ Mục 1.20 Tài khoản quốc gia 1993 của Liên Hợp Quốc.

- Sản phẩm không do một đơn vị thể chế nào tạo ra như: Phát triển tự nhiên của rừng cây, đàn cá ở sông, biển... không thuộc phạm trù sản xuất.

- Sản phẩm là hàng hóa và dịch vụ phải có khả năng cung cấp cho một đơn vị thể chế khác cho dù có thu tiền hay không. Tiêu chuẩn này của khái niệm sản xuất nhằm loại trừ các hoạt động tạo ra dịch vụ để tự tiêu dùng trong nội bộ hộ gia đình như: Nuôi dạy con cái học tập, nấu nướng, chuẩn bị bữa ăn, quét dọn, sắp xếp nhà cửa.

- Khái niệm sản xuất bao gồm cả các hoạt động bất hợp pháp tạo ra hàng hóa và dịch vụ cung cấp cho thị trường như: Buôn lậu... và các hoạt động hợp pháp nhưng tạo ra các sản phẩm bất hợp pháp.

4.1.2. Đơn vị thường trú⁽²⁾

Đơn vị thể chế được gọi là đơn vị thường trú của một quốc gia nếu nó có *trung tâm lợi ích kinh tế trong lãnh thổ kinh tế* của quốc gia đó. Một đơn vị thể chế được gọi là có trung tâm lợi ích kinh tế trong lãnh thổ kinh tế của một quốc gia nếu đơn vị đó có trụ sở, có địa điểm sản xuất hoặc nhà cửa trong lãnh thổ kinh tế của quốc gia, tiến hành các hoạt động sản xuất và giao dịch kinh tế với thời gian lâu dài. Như vậy khái niệm thường trú trong Tài khoản quốc gia không dựa trên tiêu chuẩn quốc tịch hay tiêu chuẩn pháp lý của quốc gia. Tiêu thức về trụ sở đơn vị, địa điểm sản xuất liên quan tới đơn vị sản xuất trong khi đó tiêu thức về nhà cửa liên quan tới hộ gia đình và các thành viên của hộ gia đình.

Lãnh thổ kinh tế của một quốc gia bao gồm lãnh thổ địa lý chịu sự quản lý của Nhà nước mà ở đó cư dân, hàng hóa, tài sản và vốn được tự do lưu thông. Những quốc gia có biển, lãnh thổ kinh tế còn bao gồm các hòn đảo thuộc quốc gia đó và chịu sự điều chỉnh của những chính sách tài khóa và tiền tệ như đất liền. Cụ thể, lãnh thổ kinh

⁽²⁾ Mục 3.6 đến 3.8 trong cuốn Phương pháp biên soạn hệ thống tài khoản quốc gia ở Việt Nam, Nhà xuất bản Thống kê, Hà Nội - 2003.

tế của một quốc gia bao gồm:

- Vùng đất, vùng trời, thềm lục địa nằm trong lãnh hải quốc tế mà quốc gia có quyền bất khả xâm phạm trong khai thác cá và các tài nguyên;

- Lãnh thổ quốc gia ở nước ngoài sử dụng cho mục đích ngoại giao (đại sứ quán, lãnh sự quán), mục đích quân sự (căn cứ quân sự), nghiên cứu khoa học (trạm nghiên cứu khoa học)...

4.1.3. Đơn vị thể chế⁽¹⁾

Đơn vị thể chế là đơn vị thống kê tổng quát nhất và được định nghĩa như sau: "Đơn vị thể chế là một thực thể kinh tế có quyền sở hữu tích sản, phát sinh tiêu sản và thực hiện các hoạt động, các giao dịch kinh tế với những thực thể kinh tế khác". Đơn vị thể chế có các thuộc tính sau:

- Có quyền sở hữu hàng hóa và tài sản, do vậy đơn vị thể chế có thể trao đổi quyền sở hữu này thông qua hoạt động giao dịch với đơn vị thể chế khác;

- Có trách nhiệm và chịu trách nhiệm trước pháp luật đối với những quyết định kinh tế của mình và đối với các hoạt động kinh tế có liên quan của đơn vị;

- Có khả năng phát sinh tiêu sản (có quyền huy động vốn), thực hiện các nghĩa vụ, cam kết và có tư cách pháp nhân tham gia vào các hợp đồng;

- Có điều kiện lập các tài khoản trong hệ thống tài khoản kế toán, trong đó có cả bảng cân đối kế toán theo yêu cầu của quản lý sản xuất và pháp luật của Nhà nước.

Trong thực tế, đơn vị thể chế được chia làm hai loại: Đơn vị thể chế hộ gia đình (gồm một người hay một nhóm người hình thành hộ)

⁽¹⁾ Mục 3.32 đến 3.34 sách đã dẫn.

và tổ chức kinh tế, chính trị, xã hội được pháp luật thừa nhận. Ở Việt Nam, đơn vị thể chế bao gồm các loại: Hộ gia đình tiêu dùng và hộ sản xuất kinh doanh cá thể; doanh nghiệp thuộc mọi thành phần kinh tế; cơ quan hành chính và sự nghiệp; tổ chức chính trị, chính trị - xã hội; tổ chức không vì lợi.

4.1.4. Giá cơ bản, giá sản xuất và giá sử dụng⁽¹⁾

Thống kê tài khoản quốc gia dùng ba loại giá để xác định giá trị của các chỉ tiêu kinh tế tổng hợp:

- *Giá cơ bản* là số tiền người sản xuất nhận được do bán một đơn vị hàng hoá hay dịch vụ sản xuất ra, trừ đi toàn bộ thuế đánh vào sản phẩm, cộng với trợ cấp sản phẩm. Giá cơ bản loại trừ phí vận tải không do người sản xuất trả khi bán hàng.
- *Giá sản xuất* là số tiền người sản xuất nhận được do bán một đơn vị hàng hoá hay dịch vụ sản xuất ra, trừ đi thuế giá trị gia tăng (VAT) hay thuế được khấu trừ tương tự. Giá sản xuất loại trừ phí vận tải không do người sản xuất trả khi bán hàng.
- *Giá sử dụng* là số tiền người mua phải trả để nhận được một đơn vị hàng hóa hay dịch vụ tại thời gian và địa điểm do người mua yêu cầu. Giá sử dụng không bao gồm thuế giá trị gia tăng được khấu trừ hay thuế tương tự được khấu trừ. Giá sử dụng bao gồm cả phí vận tải do người mua phải trả.

Ba loại giá trên có mối liên hệ sau:

- Giá sản xuất bằng giá cơ bản cộng với thuế sản phẩm nhưng không bao gồm thuế giá trị gia tăng (VAT), hay thuế được khấu trừ tương tự do người mua phải trả và trừ đi trợ cấp sản phẩm;
- Giá sử dụng bằng giá sản xuất cộng với thuế VAT không được khấu trừ hay loại thuế tương tự không được khấu trừ, cộng với phí vận

⁽¹⁾ Một số thuật ngữ thống kê thông dụng, Nhà xuất bản Thống kê, Hà Nội - 2004, trang 85; mục 3.65 và 3.66.

tải và phí thương nghiệp do đơn vị khác cung cấp;

- Trường hợp người sử dụng mua trực tiếp từ người sản xuất (không qua thương nghiệp bán buôn hay bán lẻ), giá sử dụng lớn hơn giá sản xuất do hai yếu tố sau: Giá trị của thuế giá trị gia tăng không được khấu trừ do người mua phải nộp; phí vận tải do người mua phải trả khi mua hàng hóa.

Mối liên hệ giữa ba loại giá nêu trên được mô tả qua sơ đồ sau:

Giá cơ bản	Thuế SP (không gồm VAT) trừ trợ cấp SP	
Giá sản xuất		Thuế VAT, Phí vận tải, Phí thương nghiệp
Giá sử dụng		

Giá sản xuất là giá "ngoại lai" giữa giá cơ bản và giá sử dụng vì nó không bao gồm một số loại thuế sản phẩm. Giá sản xuất không phải là số tiền người sản xuất thực sự nhận được khi bán sản phẩm và cũng không phải số tiền người sử dụng thực sự phải trả khi mua hàng. Nhà sản xuất dựa vào giá cơ bản để đưa ra các quyết định kinh tế; trong khi đó người tiêu dùng dựa vào giá sử dụng để quyết định việc mua hàng.

- *Giá thị trường* là giá thực tế thoả thuận giữa các thực thể kinh tế khi thực hiện các hoạt động giao dịch. Trong nền kinh tế áp dụng hệ thống thuế được khấu trừ như thuế VAT sẽ dẫn tới hai loại giá thực tế thoả thuận cho một hoạt động giao dịch nếu đứng trên quan điểm của nhà sản xuất (giá cơ bản) và người sử dụng (giá sử dụng).

- *Giá thực tế* là giá dùng trong giao dịch của năm báo cáo. Giá

thực tế phản ánh giá trị trên thị trường của hàng hóa, dịch vụ, tài sản chu chuyển từ quá trình sản xuất, lưu thông phân phối tới sử dụng cuối cùng đồng thời với sự vận động của tiền tệ, tài chính và thanh toán. Qua đó giúp ta nhận thức đúng đắn thực tiễn khách quan về cơ cấu kinh tế, mối quan hệ tỷ lệ giữa các ngành trong sản xuất, mối quan hệ phân phối thu nhập, mối quan hệ giữa kết quả sản xuất với phân huy động được vào ngân sách... trong từng năm.

- *Giá so sánh* là giá thực tế của năm được chọn làm gốc để so sánh. Để nghiên cứu sự thay đổi đơn thuần về mặt khối lượng, tức là loại trừ sự biến động của yếu tố giá, các chỉ tiêu kinh tế tổng hợp của những năm khác nhau được tính theo giá của năm gốc. Năm được chọn làm gốc để tính theo giá so sánh thường là năm trước của năm báo cáo hoặc năm đầu của thời kỳ kế hoạch.

4.1.5. Thu nhập sở hữu

Thu nhập sở hữu là thu nhập nhận được của người sở hữu tài sản tài chính hoặc tài sản hữu hình phi tài chính không do sản xuất tạo ra (như đất đai, vùng trời, vùng biển, v.v...) khi họ cung cấp tài chính hoặc đưa tài sản hữu hình phi tài chính không do sản xuất tạo ra cho đơn vị khác sử dụng. Thu nhập sở hữu bao gồm các loại sau:

- Lãi tiền gửi, tiền cho vay
- Cổ tức, lợi tức đầu tư vào các tài sản tài chính
- Thu nhập từ tái đầu tư của đầu tư trực tiếp nước ngoài;
- Thu từ cho thuê đất đai, vùng trời, vùng biển.

Thu nhập sở hữu thuần từ bên ngoài phản ánh chênh lệch về thu nhập sở hữu của một quốc gia với bên ngoài. Nếu thu nhập sở hữu thuần từ bên ngoài là dương sẽ làm tăng tổng thu nhập quốc gia, ngược lại nếu thu nhập sở hữu thuần từ bên ngoài là âm sẽ làm giảm tổng thu nhập quốc gia.

4.1.6. Chuyển nhượng

Chuyển nhượng là hoạt động giao dịch khi một đơn vị thể chế cung cấp hàng hóa, dịch vụ hoặc tài sản cho một đơn vị thể chế khác mà không nhận lại tiền, hàng hóa, dịch vụ hoặc tài sản tương ứng. Chuyển nhượng có thể bằng tiền hoặc hiện vật. Chuyển nhượng bằng tiền có thể dưới dạng tiền mặt hoặc tiền ký gửi có khả năng chuyển nhượng. Chuyển nhượng bằng hiện vật là hoạt động chuyển quyền sở hữu hàng hóa hay tài sản vật chất hoặc cung cấp dịch vụ. Theo mục đích, chuyển nhượng được chia thành hai nhóm:

Chuyển nhượng hiện hành là trao đổi thu nhập giữa các đối tượng giao dịch, làm giảm thu nhập của đơn vị thể chế cho và làm tăng thu nhập của đơn vị thể chế nhận, với mục đích *để chi tiêu dùng cuối cùng*.

Chuyển nhượng tài sản thực hiện giữa đơn vị thể chế này với mục đích cung cấp tài sản hoặc tài chính cho đơn vị thể chế kia để tích lũy tài sản.

4.1.7. Biến điểm và biến kỳ

- *Biến điểm* là khái niệm biểu thị giá trị tại một thời điểm nhất định, dùng để đánh giá các chỉ tiêu kinh tế tổng hợp tại các thời điểm đầu kỳ và cuối kỳ như: Giá trị tài sản cố định và tài sản lưu động, giá trị của cải của toàn bộ nền kinh tế, v.v... Trong hệ thống tài khoản quốc gia, khái niệm biến điểm được áp dụng khi biên soạn bảng tổng kết tài sản.

- *Biến kỳ* là khái niệm giá trị trong một khoảng thời gian, dùng để đánh giá các chỉ tiêu tổng hợp của nền kinh tế như: giá trị sản xuất, tổng sản phẩm trong nước, tích lũy, tiêu dùng, tổng thu nhập quốc gia, để dành, v.v... Chẳng hạn chỉ tiêu tổng sản phẩm trong nước áp dụng khái niệm biến kỳ với nghĩa đánh giá giá trị hàng hóa và dịch vụ cuối cùng được tạo ra của nền kinh tế trong một khoảng thời gian như quý hoặc năm.

Khái niệm biến điểm và biến kỳ có mối quan hệ với nhau, biến

điểm là kết quả của các hoạt động giao dịch dồn tích lại trong một thời kỳ nhất định. Chẳng hạn, giá trị tài sản cố định tại thời điểm 31 tháng 12 là kết quả của những hoạt động giao dịch về tài sản cố định diễn ra từ 1 tháng 1 đến 30 tháng 12.

4.1.8. Tích sản và tiêu sản

- *Tích sản* là một thực thể có chức năng lưu giữ giá trị, qua đó các đơn vị thể chế trong nền kinh tế xác lập quyền sở hữu đối với nó và thu được lợi ích kinh tế qua việc sở hữu, sử dụng theo thời gian.

- *Tiêu sản* phản ánh bốn phận hay trách nhiệm của một đơn vị thể chế phải thanh toán cho một đơn vị thể chế khác trong những trường hợp được quy định cụ thể theo hợp đồng giữa hai đơn vị có liên quan.

Trong kinh tế, tích sản được chia thành hai loại: Tích sản tài chính và tích sản phi tài chính. Tích sản tài chính bao gồm trái quyền tài chính, vàng, tiền, quyền rút vốn đặc biệt (SDR), cổ phiếu công ty và các công cụ tài chính kinh doanh ngoài bảng. Tích sản phi tài chính là tích sản không có yếu tố tiêu sản tương ứng, bao gồm các tài sản vật chất hữu hình, vô hình và tài sản không do sản xuất tạo ra như đất đai.

Các nhà kinh tế thiết lập các công cụ tài chính kinh doanh ngoài bảng với mục đích tránh cho các bên có liên quan trong giao dịch khỏi chịu thiệt trong tương lai khi giá cả biến động lớn. Những công cụ kinh doanh ngoài bảng gồm: hợp đồng mua bán trước; giao dịch có kỳ hạn; giao dịch hoán đổi. Hợp đồng mua bán trước cho phép một bên được mua hoặc bán hàng hóa hay chứng khoán trong một thời hạn nhất định với mức giá thỏa thuận trước. Đây là biểu hiện của hình thức đầu cơ vì nếu giá cả thay đổi một cách đáng kể thì người mua vẫn được mua với mức giá thỏa thuận trước mà chắc chắn thấp hơn nhiều so với mức giá hiện thời. Giao dịch có kỳ hạn thường liên quan tới thị trường ngoại hối, ở đó các đồng tiền được mua và bán theo những tỷ giá hối đoái được cố định tại thời điểm

mua và giao vào một thời gian nhất định trong tương lai. Giao dịch hoán đổi là phương pháp hoán đổi các đồng tiền. Ngân hàng trung ương của hai nước ghi Có cho nhau với một khoản tiền của họ có giá trị tương đương để mỗi chính phủ đều có thể sử dụng dự trữ ngoại hối này nếu cần thiết.

4.1.9. Chỉ tiêu cân đối

Chỉ tiêu cân đối trong các tài khoản của Hệ thống tài khoản quốc gia là chỉ tiêu được thiết lập theo nguyên tắc cân bằng của tài khoản, thu được bằng cách lấy tổng bên nguồn trừ đi tất cả các chỉ tiêu bên sử dụng. Chỉ tiêu cân đối không liên quan tới bất kỳ một tập hợp các giao dịch cụ thể nào và cũng không tính theo một đơn vị giá cả hay đơn vị khối lượng cụ thể nào. Nói cách khác, chỉ tiêu cân đối không thể phân tích thành hai yếu tố giá và lượng vì vậy không thể tính trực tiếp các chỉ tiêu này theo giá so sánh. Các nhà Thống kê cần lưu ý tới đặc trưng này của chỉ tiêu cân đối trong khi tính toán.

Đưa ra chỉ tiêu cân đối trong Hệ thống tài khoản quốc gia không chỉ đơn giản nhằm mục đích làm cân bằng giữa bên nguồn và bên sử dụng của các tài khoản. Chỉ tiêu cân đối đã chứa đựng một khối lượng lớn thông tin hữu ích trong nó và bao gồm một số chỉ tiêu đầu vào quan trọng nhất trong các tài khoản của hệ thống như các chỉ tiêu: Giá trị tăng thêm, thu nhập khả dụng, để dành, cho vay thuần hoặc đi vay thuần⁽¹⁾.

4.2. MỘT SỐ CHỈ TIÊU CHỦ YẾU TRONG THỐNG KÊ TÀI KHOẢN QUỐC GIA

Phần này sẽ trình bày tuần tự các chỉ tiêu kinh tế tổng hợp phản ánh kết quả từ quá trình sản xuất tạo ra thu nhập, phân phối thu nhập, đến tiêu dùng, để dành, tích lũy tài sản và của cải của nền kinh tế. Qua đây bạn đọc thấy được một cách khái quát nội dung, bản chất kinh tế

⁽¹⁾ Mục 3.65, Tài khoản quốc gia 1993.

và mối liên hệ giữa những chỉ tiêu kinh tế này. Trước hết, chúng ta nghiên cứu chỉ tiêu đầu tiên xuất hiện bên nguồn của tài khoản sản xuất đó là giá trị sản xuất.

4.2.1. Giá trị sản xuất

Giá trị sản xuất phản ánh toàn bộ giá trị của sản phẩm vật chất (thành phẩm, bán thành phẩm, sản phẩm dở dang) và dịch vụ sản xuất ra trong một thời kỳ nhất định. Giá trị sản xuất bao gồm:

- Giá trị hàng hóa và dịch vụ sử dụng hết trong quá trình sản xuất;
- Giá trị mới tăng thêm trong quá trình sản xuất: Thu nhập của người lao động từ sản xuất, thuế sản xuất, khấu hao tài sản cố định dùng trong sản xuất và thặng dư sản xuất.

Giá trị sản xuất được tính theo các ngành kinh tế và tổng hợp cho toàn bộ nền kinh tế. Giá trị sản xuất có sự tính trùng giữa các đơn vị trong từng ngành cũng như giữa các ngành trong nền kinh tế quốc dân. Mức độ tính trùng phụ thuộc vào trình độ chuyên môn hóa trong sản xuất của nền kinh tế. Trình độ chuyên môn hóa càng cao thì sự tính trùng càng nhiều.

Giá trị sản xuất được xác định theo giá cơ bản; khi không có điều kiện về nguồn thông tin, chế độ hạch toán và kế toán không phù hợp thì có thể tính theo giá sản xuất.

Để tạo ra hàng hóa và dịch vụ đáp ứng nhu cầu của nền kinh tế, các nhà sản xuất phải sử dụng lao động và máy móc thiết bị để chuyển những chi phí là vật chất và dịch vụ thành hàng hóa và dịch vụ mới. Trong kinh tế vĩ mô cũng như thống kê tài khoản quốc gia, giá trị hàng hóa và dịch vụ chuyển hóa trong quá trình sản xuất được gọi là chi phí trung gian. Nói cách khác, *Chi phí trung gian* phản ánh giá trị hàng hóa và dịch vụ được sử dụng hết trong quá trình sản xuất để tạo ra sản phẩm mới trong một thời kỳ nhất định, gồm cả chi phí sửa chữa nhỏ và duy tu tài sản cố định dùng trong sản xuất. Chi phí trung gian được tính theo ngành kinh tế và toàn bộ nền kinh tế, theo giá thực tế

và giá so sánh. Chi phí trung gian chia thành hai nhóm chủ yếu:

- Nhóm chi phí vật chất gồm: Nguyên vật liệu chính, vật liệu phụ, nhiên liệu, điện, nước, khí đốt, chi phí công cụ sản xuất nhỏ, vật rẻ tiền mau hỏng và chi phí sản phẩm vật chất khác;

- Nhóm chi phí dịch vụ gồm: Vận tải; bưu điện; bảo hiểm; dịch vụ ngân hàng; dịch vụ pháp lý, dịch vụ quảng cáo và các dịch vụ khác.

Khái niệm chi phí trung gian trong kinh tế vĩ mô cũng như thống kê tài khoản quốc gia khác với khái niệm chi phí sản xuất trong kinh tế vi mô. Có những loại chi tiêu, kinh tế vi mô tính vào chi phí sản xuất như: Chi trả lương cho người lao động; chi ăn trưa, ca ba bằng tiền cho người lao động; chi mua sắm quần áo cho người lao động, v.v... nhưng thống kê tài khoản quốc gia lại tính vào thu nhập của người lao động và thuộc giá trị tăng thêm, không thuộc chi phí trung gian. Chi phí sản xuất trong kinh tế vi mô còn gồm cả khấu hao tài sản cố định, nhưng tài khoản quốc gia lại đưa khấu hao tài sản cố định vào giá trị tăng thêm. Vì vậy, không thể đồng nhất khái niệm chi phí trung gian với khái niệm chi phí sản xuất.

Không giống như giá trị sản xuất có thể tính theo giá cơ bản hoặc giá sản xuất, chi tiêu chi phí trung gian chỉ tính theo giá sử dụng. Giá sử dụng dùng để xác định giá trị của chi phí trung gian phản ánh toàn bộ số tiền đơn vị sản xuất phải trả để đưa một đơn vị hàng hóa hoặc dịch vụ tham gia vào quá trình sản xuất với mục đích tạo ra hàng hóa và dịch vụ mới. Giá sử dụng dùng để tính chi phí trung gian gồm hai phần chính: Giá trị hàng hóa đơn vị sản xuất mua từ đơn vị bán, phí vận tải phát sinh để chuyên chở hàng hóa từ nơi bán tới đơn vị sản xuất.

Như đã đề cập ở trên, chỉ tiêu giá trị sản xuất có sự tính trùng trong nền kinh tế, giá trị sản xuất bao gồm cả giá trị hàng hóa và dịch vụ được tạo ra của các thời kỳ sản xuất trước, chẳng hạn dùng nguyên vật liệu được tạo ra của năm trước để sản xuất ra hàng hóa và dịch vụ năm sau. Với đặc điểm đó, nếu dùng giá trị sản xuất để đánh giá tốc

độ tăng trưởng kinh tế và các chỉ tiêu tổng hợp liên quan tới bình quân đầu người, năng suất, hiệu quả thì sẽ không phản ánh đúng kết quả sản xuất của nền kinh tế trong một thời kỳ nhất định. Để khắc phục tồn tại này, các nhà kinh tế và thống kê đã tính toán và đưa vào áp dụng chỉ tiêu giá trị tăng thêm và tổng sản phẩm trong nước.

4.2.2. Giá trị tăng thêm

Giá trị tăng thêm (VA) phản ánh giá trị hàng hóa và dịch vụ cuối cùng được tạo ra của các *ngành kinh tế* trong một thời kỳ nhất định. Giá trị tăng thêm là bộ phận còn lại của giá trị sản xuất sau khi trừ đi phần chi phí trung gian. Cấu thành của giá trị tăng thêm gồm: Thu nhập của người lao động từ sản xuất, thuế sản xuất, khấu hao tài sản cố định dùng trong sản xuất và thặng dư sản xuất.

Cũng như giá trị sản xuất, giá trị tăng thêm có thể tính theo giá cơ bản hoặc giá sản xuất. Giá trị sản xuất tính theo giá nào đòi hỏi giá trị tăng thêm tính theo giá đó. Nói cách khác, giá trị sản xuất và giá trị tăng thêm luôn tính theo cùng một loại giá.

Vì không bao gồm chi phí trung gian, nên giá trị tăng thêm phản ánh đúng mức hơn về kết quả hoạt động sản xuất do đơn vị tạo ra, không bị phụ thuộc vào thay đổi tổ chức sản xuất của đơn vị. Do vậy, dùng chỉ tiêu giá trị tăng thêm để đánh giá kết quả sản xuất của từng ngành, từng doanh nghiệp cũng như tính các chỉ tiêu năng suất, hiệu quả và nhiều chỉ tiêu quan trọng khác có liên quan sẽ có ý nghĩa hơn, phản ánh thực chất hơn thành quả lao động của đơn vị và ngành so với áp dụng chỉ tiêu giá trị sản xuất như đã nói ở trên.

Giá trị tăng thêm được tính theo ngành kinh tế. Để phản ánh giá trị hàng hóa và dịch vụ cuối cùng được tạo ra của toàn bộ nền kinh tế, các nhà kinh tế vĩ mô đưa ra khái niệm tổng sản phẩm trong nước.

4.2.3. Tổng sản phẩm trong nước

Tổng sản phẩm trong nước (GDP) phản ánh giá trị của hàng hóa

và dịch vụ cuối cùng được tạo ra của *toàn bộ nền kinh tế* trong một khoảng thời gian nhất định. Tổng sản phẩm trong nước là khái niệm của giá trị tăng thêm - khái niệm trên giác độ sản xuất và bằng tổng giá trị tăng thêm của tất cả các đơn vị sản xuất thường trú trong lãnh thổ kinh tế của một quốc gia cộng với thuế nhập khẩu hàng hóa và dịch vụ.

Dựa vào phân tích luồng chu chuyển thu nhập và chi tiêu trong nền kinh tế, các nhà kinh tế đã chỉ ra mối quan hệ đẳng thức giữa tổng sản phẩm trong nước với tổng thu nhập từ sản xuất và tổng chi tiêu trong nền kinh tế. Các đẳng thức này là cơ sở lý luận của ba phương pháp tính chỉ tiêu tổng sản phẩm trong nước: Phương pháp sản xuất bằng tổng giá trị tăng thêm của các ngành kinh tế cộng (+) thuế nhập khẩu hàng hóa và dịch vụ; phương pháp thu nhập bằng tổng các yếu tố (thu nhập của người lao động từ sản xuất, thuế sản xuất, khấu hao TSCĐ, thặng dư) và phương pháp sử dụng bằng tổng của tiêu dùng cuối cùng, tích lũy tài sản, chênh lệch xuất nhập khẩu hàng hóa và dịch vụ.

Tổng sản phẩm trong nước luôn tính theo giá thị trường, nói cách khác chỉ có một loại giá dùng để tính GDP mặc dù chỉ tiêu giá trị tăng thêm có thể tính theo giá cơ bản hay giá sản xuất. Có thể đặt câu hỏi tại sao giá trị tăng thêm của các ngành kinh tế tính theo giá cơ bản hoặc giá sản xuất nhưng GDP chỉ tính theo giá thị trường? Phải chăng có sự không thống nhất về giá giữa các phương pháp tính GDP? Giá trị tăng thêm của các ngành sản xuất vật chất và dịch vụ (không bao gồm ngành thương nghiệp và vận tải) theo giá cơ bản hoặc giá sản xuất không bao gồm phí vận tải và phí thương nghiệp. Khi cộng giá trị tăng thêm của tất cả các ngành kinh tế theo giá cơ bản hoặc giá sản xuất ta được tổng sản phẩm trong nước đã bao gồm giá trị phí thương nghiệp và phí vận tải.

Công thức chung tính tổng sản phẩm trong nước đối với trường hợp giá trị tăng thêm tính theo giá cơ bản và giá sản xuất lần lượt như sau:

$$\begin{array}{ccccccc} \text{Tổng} & & \text{Tổng giá trị} & & \text{Thuế sản} & & \text{Thuế nhập} \\ \text{sản phẩm} & = & \text{tăng thêm theo} & + & \text{phẩm trừ trợ} & + & \text{khẩu hàng hóa} \\ \text{trong nước} & & \text{giá cơ bản} & & \text{cấp sản phẩm} & & \text{và dịch vụ} \end{array}$$

Hoặc

$$\begin{array}{ccccccc} \text{Tổng sản} & & \text{Tổng giá trị} & & \text{Thuế nhập} & & \\ \text{phẩm trong} & = & \text{tăng thêm theo} & + & \text{khẩu hàng hóa} & & \\ \text{nước} & & \text{giá sản xuất} & & \text{và dịch vụ} & & \end{array}$$

Theo phương pháp sử dụng, GDP bằng tổng sử dụng hàng hóa và dịch vụ bao gồm: Tiêu dùng cuối cùng của hộ gia đình và Nhà nước, tích lũy tài sản cố định, tài sản lưu động và tài sản quý hiếm, chênh lệch giữa xuất khẩu và nhập khẩu hàng hóa và dịch vụ. Giá sử dụng dùng để xác định giá trị của tất cả các yếu tố trong tổng sử dụng hàng hóa và dịch vụ.

Trên góc độ toàn bộ nền kinh tế dùng chỉ tiêu GDP để đánh giá tốc độ tăng trưởng kinh tế cũng như tính toán các chỉ tiêu liên quan sẽ có những ưu điểm tương tự như áp dụng chỉ tiêu giá trị tăng thêm trong từng ngành, từng đơn vị.

4.2.4. Tiêu dùng cuối cùng

Tiêu dùng cuối cùng phản ánh toàn bộ chi tiêu cho mua sắm hàng hóa và dịch vụ tiêu dùng của hộ gia đình thường trú, của các tổ chức không vì lợi phục vụ hộ gia đình thường trú và của Nhà nước trong một thời kỳ nhất định. Tiêu dùng cuối cùng là một bộ phận của thu nhập quốc gia khả dụng và cũng là một bộ phận của tổng sản phẩm trong nước.

Tiêu dùng cuối cùng được chia theo nhóm hàng hóa và dịch vụ tiêu dùng, được tính theo giá thực tế và giá so sánh và thường tách thành hai thành phần: Tiêu dùng cuối cùng của hộ gia đình và tiêu dùng cuối cùng của Nhà nước.

4.2.5. Tích lũy tài sản

Tích lũy tài sản phản ánh chi tiêu cho đầu tư tài sản cố định, đầu tư tài sản lưu động và tài sản quý hiếm trong một thời kỳ nhất định. Tích lũy tài sản được chia theo loại tài sản, tính theo giá thực tế và giá so sánh.

+ Tích lũy tài sản cố định tính bằng giá trị tài sản cố định nhận về trừ đi tài sản cố định thanh lý trong kỳ của các đơn vị thể chế, không bao gồm phần hộ gia đình tiêu dùng.

+ Tích lũy tài sản lưu động gồm tài sản là nguyên, nhiên, vật liệu dùng cho sản xuất, thành phẩm tồn kho, sản phẩm dở dang; được tính bằng chênh lệch giữa tài sản lưu động nhận được và tài sản lưu động sử dụng trong kỳ của các đơn vị thể chế, không bao gồm phần hộ gia đình tiêu dùng.

+ Tài sản quý hiếm do các đơn vị thể chế (gồm cả hộ gia đình tiêu dùng) nắm giữ với mục đích bảo toàn giá trị của cải. Tài sản quý hiếm không bị hao mòn và giảm giá trị theo thời gian, được tính bằng chênh lệch giữa tài sản quý hiếm nhận được trong kỳ và nhượng bán tài sản quý hiếm trong kỳ đó.

4.2.6. Xuất, nhập khẩu hàng hóa và dịch vụ

Xuất, nhập khẩu hàng hóa và dịch vụ phản ánh toàn bộ sản phẩm vật chất và dịch vụ được mua bán, trao đổi, chuyển nhượng giữa các đơn vị, tổ chức, cá nhân dân cư là đơn vị thường trú của Việt Nam với các đơn vị, tổ chức, cá nhân dân cư không thường trú. Những hàng hóa và dịch vụ được coi là xuất, nhập khẩu khi đã thay đổi quyền sở hữu về hàng hóa giữa một bên là đơn vị thường trú và bên kia là đơn vị không thường trú, không phụ thuộc vào hàng hóa đó đã ra khỏi biên giới quốc gia hay chưa.

Chênh lệch xuất nhập khẩu phản ánh hiệu số giữa trị giá xuất khẩu hàng hóa và dịch vụ với trị giá nhập khẩu hàng hóa và dịch vụ của một quốc gia trong một thời kỳ nhất định. Nếu hiệu số này là một

số dương thì nền kinh tế được gọi là có xuất siêu, ngược lại nếu hiệu số này là một số âm thì nền kinh tế được gọi là nhập siêu. Trường hợp trị giá xuất khẩu hàng hóa và dịch vụ bằng với trị giá nhập khẩu hàng hóa và dịch vụ, nền kinh tế có cân bằng trong quan hệ thương mại với bên ngoài.

Như đã đề cập trong phần khái niệm, sản xuất là quá trình sử dụng lao động và máy móc thiết bị để chuyển những chi phí là vật chất và dịch vụ thành sản phẩm là vật chất và dịch vụ khác. Vì vậy, tổng sản phẩm trong nước tính theo phương pháp thu nhập sẽ bằng tổng của các yếu tố tham gia vào quá trình sản xuất. Dưới dạng giá trị, những yếu tố này gồm: Thu nhập của người lao động từ sản xuất; thuế, trợ cấp sản xuất; khấu hao tài sản cố định và thặng dư.

4.2.7. Thu nhập của người lao động từ sản xuất

Thu nhập của người lao động từ sản xuất phản ánh tổng thu nhập của người lao động từ tất cả các nguồn do tham gia vào quá trình sản xuất mang lại, gồm thu nhập bằng tiền và bằng hiện vật. Cụ thể thu nhập của người lao động bao gồm: Tiền lương lĩnh đều đặn theo kỳ; tiền lương trả cho người lao động do nghỉ việc tạm thời vì lý do của đơn vị sản xuất; tiền hoa hồng do đơn vị sản xuất trả cho người lao động; tiền thưởng đột xuất và các khoản thanh toán cho người lao động liên quan tới hoạt động sản xuất; chi hỗ trợ định kỳ về nhà ở và đi lại từ nơi ở tới nơi làm việc; tiền lương bằng hiện vật; tiền đóng bảo hiểm xã hội cho người lao động.

4.2.8. Thuế sản xuất, trợ cấp sản xuất

Thuế sản xuất, trợ cấp sản xuất là yếu tố thứ hai cấu thành nên GDP tính theo phương pháp thu nhập. Thuế sản xuất là khoản phải nộp bắt buộc bằng tiền hay bằng hiện vật từ đơn vị sản xuất cho nhà nước khi tham gia vào quá trình sản xuất, lưu thông hàng hóa và dịch vụ. Thuế sản xuất gồm hai loại: Thuế sản phẩm và thuế sản xuất khác:

- Thuế sản phẩm phải nộp khi người sản xuất đưa hàng hóa và

dịch vụ vào lưu thông dưới bất kỳ hình thức nào như: Bán, chuyển nhượng... Như vậy, đối tượng của thuế đánh vào sản phẩm không bao gồm thành phẩm tồn kho. Loại thuế này gồm cả thuế hàng nhập khẩu khi hàng nhập khẩu đi vào lãnh thổ kinh tế hay dịch vụ phục vụ cho đơn vị thường trú từ đơn vị không thường trú;

- Thuế sản xuất khác bao gồm thuế đánh vào quyền sở hữu hay quyền sử dụng đất đai, nhà xưởng, tài sản khác dùng trong sản xuất; hay thuế đánh vào thuê mướn lao động, trả thu nhập cho người lao động⁽¹⁾.

4.2.9. Khấu hao tài sản cố định

Khấu hao tài sản cố định phản ánh giá trị của tài sản cố định tiêu dùng trong quá trình sản xuất và tính bằng chênh lệch giữa cuối kỳ và đầu kỳ giá trị kinh tế thực của tài sản (giá trị kinh tế thực của tài sản là giá trị thực tế của tài sản trên thị trường tại thời điểm đánh giá). Trong thực tế, tính khấu hao tài sản cố định dựa trên cơ sở thời gian dự kiến dùng vào sản xuất của tài sản và giá trị tài sản theo nguyên giá.

4.2.10. Thặng dư

Thặng dư là thu nhập từ sản xuất của đơn vị sản xuất và được tính bằng giá trị tăng thêm trừ đi thu nhập của người lao động từ sản xuất, trừ thuế sản xuất phải nộp và cộng với trợ cấp sản xuất. Thặng dư biểu thị thu nhập có được từ quá trình sản xuất đưa lại trước khi chi trả lãi tiền vay ngân hàng, tiền thuê máy móc thiết bị, thu nhập sở hữu phải trả đối với tài sản tài chính, tiền thuê đất cần thiết để tiến hành sản xuất.

* *
*

Ba phương pháp tính chỉ tiêu GDP được xây dựng trên ba góc độ

⁽¹⁾ Mục 3.46 Phương pháp biên soạn Hệ thống tài khoản quốc gia ở Việt Nam, Nhà xuất bản Thống kê, Hà Nội, 2003.

khác nhau: Phương pháp sản xuất thực hiện trên góc độ sản xuất tạo ra sản phẩm là hàng hóa và dịch vụ cho xã hội; phương pháp thu nhập đứng trên góc độ các yếu tố tham gia vào quá trình sản xuất tạo ra thu nhập; phương pháp sử dụng đứng trên góc độ sử dụng hàng hóa và dịch vụ để thỏa mãn nhu cầu cuối cùng của nền kinh tế. Mỗi phương pháp đều có ưu và nhược điểm riêng và sử dụng phương pháp nào trong tính toán chỉ tiêu GDP phụ thuộc vào nguồn thông tin hiện có, trình độ thống kê và điều kiện hạch toán trong từng thời kỳ khác nhau.

Phương pháp sản xuất: Tổng sản phẩm trong nước tính theo phương pháp này thường được áp dụng tại các nước đang phát triển, có trình độ thống kê chưa cao. Áp dụng phương pháp sản xuất cung cấp cho các nhà quản lý, lập chính sách bức tranh toàn cảnh về toàn bộ sản phẩm sản xuất ra, về chi phí của tất cả các ngành kinh tế. Trên cơ sở đó, các nhà kinh tế có thể nghiên cứu tác động qua lại giữa các ngành, thành phần kinh tế và tìm ra các nguyên nhân cũng như giải pháp để giảm tỷ lệ của chi phí trung gian dẫn tới tăng GDP. Bên cạnh những ưu điểm nêu trên, phương pháp sản xuất có một số nhược điểm chủ yếu sau:

- Khó đảm bảo phạm vi thu thập thông tin để tính đầy đủ kết quả của các hoạt động sản xuất. Chẳng hạn, theo khái niệm sản xuất, giá trị của các hoạt động bất hợp pháp tạo ra hàng hóa và dịch vụ cung cấp cho thị trường và các hoạt động hợp pháp nhưng tạo ra các sản phẩm bất hợp pháp đều phải tính vào giá trị sản xuất⁽¹⁾. Nhưng trong thực tế rất khó thu thập được thông tin của các hoạt động này.

- Do GDP được tính gián tiếp qua giá trị sản xuất và chi phí trung gian nên chất lượng tính GDP còn phụ thuộc vào chất lượng tính chi tiêu chi phí trung gian. Nói cách khác, thông tin để tính GDP còn phụ thuộc vào thông tin từ hạch toán chi phí sản xuất của các đơn vị sản

⁽¹⁾ Thống kê tài khoản quốc gia xếp các hoạt động này vào khu vực kinh tế chưa được quan sát. Khu vực kinh tế này gồm bốn khu vực nhỏ: Khu vực kinh tế chưa định hình, khu vực kinh tế ngầm, khu vực tự sản tự tiêu của hộ gia đình và khu vực kinh tế bất hợp pháp.

xuất kinh doanh. Xu hướng các nhà sản xuất kinh doanh thường hạch toán tăng chi phí để giảm thuế và tăng lợi nhuận. Các nhà thống kê rất khó kiểm soát chất lượng thông tin về chi phí sản xuất để từ đó tính toán chính xác, nâng cao chất lượng của chỉ tiêu chi phí trung gian.

Phương pháp sử dụng: GDP tính theo phương pháp này cung cấp những thông tin về cầu của nền kinh tế như: Tiêu dùng, tích lũy, xuất, nhập khẩu. Những thông tin này giúp cho Chính phủ và các nhà quản lý đưa ra chính sách kích cầu dẫn tới tăng trưởng kinh tế, đồng thời cũng là nguồn thông tin dùng để tính giá trị tài sản và của cải tăng thêm của đất nước. Phương pháp sử dụng có một số ưu điểm về mặt tính toán như sau:

- Phương pháp này không phải giải quyết vấn đề nan giải về việc tính giá trị của các hoạt động bất hợp pháp và các hoạt động hợp pháp nhưng tạo ra các sản phẩm bất hợp pháp như phương pháp sản xuất gặp phải. Dù không thu được thông tin về kết quả sản xuất của các hoạt động thuộc khu vực kinh tế chưa được quan sát nhưng những hoạt động này tạo ra thu nhập và thực thể kinh tế sở hữu thu nhập này sẽ dùng vào tiêu dùng, vào tích lũy và như vậy đã được phản ánh trong phương pháp sử dụng. Nếu đảm bảo được chất lượng tính toán của phương pháp sử dụng, chênh lệch về GDP giữa phương pháp sử dụng và phương pháp sản xuất là một ước lượng tốt đối với giá trị tăng thêm của khu vực kinh tế chưa được quan sát.

- Thông tin về chi tiêu dùng và tích lũy thường sát với thực tế hơn so với thông tin về kết quả sản xuất và chi phí sản xuất. Thông tin về chi tiêu dùng của hộ gia đình dễ kiểm soát và thông thường các hộ chỉ khai thấp thu nhập chứ hiếm khi khai thấp chi tiêu trong các cuộc điều tra. Thông tin về xuất, nhập khẩu hàng hóa và dịch vụ dễ thu thập (chỉ cần qua Hải quan) và thường đảm bảo về phạm vi. Vì vậy, đối với nhà thống kê, phương pháp sử dụng để tính toán và cho chất lượng số liệu cao hơn.

Tuy vậy, phương pháp sử dụng có một số nhược điểm: Số lượng hộ gia đình trong nền kinh tế rất lớn vì vậy không thể tiến hành điều tra định kỳ thường xuyên để thu thập thông tin từ tất cả các hộ. Thông thường các nước dựa vào điều tra chọn mẫu vì vậy số liệu chịu ảnh hưởng của sai số chọn mẫu. Đối với các nước đang phát triển, sản xuất nhỏ, manh mún còn phổ biến, hệ thống luật pháp chưa đầy đủ, ý thức chấp hành luật chưa nghiêm nên khó thu thập được chính xác những thông tin về tiêu dùng và xuất, nhập khẩu hàng hóa và dịch vụ. Thực tế ở nước ta, ngành Thống kê gặp không ít khó khăn khi tính toán giá trị của hàng hóa xuất, nhập khẩu lậu qua biên giới, trên biển.

Phương pháp thu nhập: GDP tính theo phương pháp này sẽ cung cấp thông tin cho các nhà quản lý và lập chính sách dùng để đánh giá về hiệu quả (thể hiện qua chỉ tiêu thặng dư) và năng suất (thể hiện qua năng suất lao động, năng suất của máy móc, thiết bị) của hoạt động sản xuất. Tuy vậy, chất lượng tính toán thấp của các chỉ tiêu cấu thành nên GDP của phương pháp này là hạn chế lớn nhất. Rất khó thu thập thông tin và tính toán chính xác thu nhập của người lao động từ sản xuất. Tâm lý của người lao động không muốn cung cấp thông tin chính xác về thu nhập của họ.

Khấu hao tài sản cố định cũng rất khó tính được chính xác vì thời gian dự kiến sử dụng trong sản xuất của tài sản thay đổi thường xuyên, giá trị của tài sản phụ thuộc vào tiến bộ của khoa học công nghệ. Máy tính điện tử là ví dụ rất điển hình minh họa cho việc khó tính toán chính xác chỉ tiêu khấu hao tài sản. Giả sử một máy tính nhãn hiệu IBM trị giá 15 triệu đồng, dự kiến sử dụng vào sản xuất trong 5 năm và khấu hao mỗi năm là 3 triệu. Do áp dụng thành quả của khoa học - kỹ thuật, máy tính mới, cùng nhãn hiệu có tính năng hoạt động cao hơn nhưng giá chỉ bằng hai phần ba giá máy của năm trước. Rõ ràng máy tính của năm trước không thể định giá là 12 triệu đồng và để nâng cao năng lực cạnh tranh, nhà sản xuất có thể rút ngắn thời gian sử dụng.

Tương tự như chỉ tiêu thu nhập của người lao động từ sản xuất, rất khó thu thập thông tin và tính chính xác được chỉ tiêu thặng dư vì đơn vị sản xuất thường hạch toán tăng chi phí để giảm thặng dư.

Từ những ưu và nhược điểm của ba phương pháp tính GDP, phương pháp sử dụng thường được áp dụng và cho kết quả tính toán với chất lượng cao nhất, tiếp đến là phương pháp sản xuất. Phương pháp thu nhập thường áp dụng khi muốn biết tỷ lệ của các nhân tố tham gia vào quá trình sản xuất. Việc quyết định áp dụng phương pháp tính nào tùy thuộc vào tình hình cụ thể của từng quốc gia và nó sẽ quyết định chiến lược thu thập thông tin đầu vào của ngành Thống kê.

Tổng sản phẩm trong nước là chỉ tiêu phản ánh hàng hóa và dịch vụ cuối cùng được tạo ra của các đơn vị thường trú trong nền kinh tế ở một thời kỳ nhất định. Nói cách khác, khái niệm tổng sản phẩm trong nước nhấn mạnh tới sản phẩm sản xuất ra trong lãnh thổ kinh tế của một quốc gia mà không quan tâm tới đặc trưng về sở hữu của các đơn vị thường trú đóng trên lãnh thổ kinh tế. Trong thực tế nhiều đơn vị thường trú trên lãnh thổ kinh tế của một quốc gia nhưng thuộc sở hữu của một quốc gia khác. Các quốc gia khác có thể sở hữu toàn bộ đơn vị sản xuất thường trú của một quốc gia như đơn vị có 100% vốn đầu tư nước ngoài, hoặc sở hữu một trong các yếu tố tham gia vào quá trình sản xuất như lao động, tài sản, vốn dùng trong sản xuất. Kết thúc quá trình sản xuất tạo ra thu nhập, chủ sở hữu của những thực thể kinh tế của các quốc gia bên ngoài có thể chuyển phần thu nhập của họ về nước. Vì vậy, không phải tất cả GDP của một quốc gia là thu nhập của quốc gia đó. Để làm rõ đặc trưng này, các nhà kinh tế vĩ mô và thống kê tài khoản quốc gia đã đưa ra khái niệm tổng thu nhập quốc gia.

4.2.11. Tổng thu nhập quốc gia

Tổng thu nhập quốc gia (GNI) phản ánh tổng thu nhập lần đầu được tạo ra từ các yếu tố thuộc sở hữu của quốc gia tham gia vào hoạt

động sản xuất trên lãnh thổ quốc gia hay ở nước ngoài trong một thời kỳ nhất định, thường là một năm.

Tổng thu nhập quốc gia bằng tổng sản phẩm trong nước cộng với chênh lệch giữa thu nhập của người lao động Việt Nam ở nước ngoài gửi về và thu nhập của người nước ngoài ở Việt Nam gửi ra bên ngoài cộng với chênh lệch giữa thu nhập sở hữu nhận được từ nước ngoài với thu nhập sở hữu trả nước ngoài.

Khái niệm GNI nhấn mạnh tới tiêu thức sở hữu các yếu tố sản xuất của một quốc gia, trong khi đó khái niệm GDP nhấn mạnh tới tiêu thức sản phẩm tạo ra trong phạm vi lãnh thổ quốc gia. Để phân biệt rõ hơn sự khác nhau giữa GDP và GNI chúng ta xét ví dụ: Giả sử công ty Honda 100% vốn của Nhật tại Việt Nam sản xuất ô tô và xe máy. Công ty này đóng trên lãnh thổ Việt Nam, nên toàn bộ giá trị sản phẩm cuối cùng do công ty tạo ra tính vào GDP của Việt Nam. Không phải toàn bộ giá trị của sản phẩm cuối cùng này được tính vào chỉ tiêu GNI của Việt Nam vì công ty thuộc sở hữu của Nhật, vì vậy phần giá trị sản phẩm cuối cùng sau khi trừ đi chi trả lương, nộp thuế sản xuất, thuế thu nhập, thuộc sở hữu của công ty Honda và được gửi về Nhật. Trong trường hợp này GNI của Việt Nam bằng GDP trừ đi phần thu nhập thuộc sở hữu của công ty Honda.

Chỉ tiêu GNI phản ánh thu nhập thực của một quốc gia do vậy trước đây các nước công nghiệp phát triển thường tính và sử dụng chỉ tiêu này. Hiện nay họ chuyển sang tính và sử dụng chỉ tiêu GDP vì:

- Phần lớn các nước đang phát triển dùng chỉ tiêu GDP vì vậy chuyển sang để thuận lợi cho so sánh quốc tế;
- Chỉ tiêu GDP dễ tính toán hơn GNI do nguồn thông tin về thu nhập sở hữu thuần rất kém;
- Chỉ tiêu GDP đánh giá tốt hơn về tiềm năng tạo việc làm của nền kinh tế so với chỉ tiêu GNI.

Cùng với tổng thu nhập quốc gia các nhà thống kê kinh tế còn đưa ra khái niệm thu nhập quốc gia thuần. Sự khác biệt giữa hai chỉ

tiêu trên là trong đó có bao gồm hay không bao gồm giá trị khấu hao tài sản cố định. Một cách tổng quát, khấu hao tài sản cố định phản ánh sự suy giảm giá trị hiện tại của tài sản cố định thuộc sở hữu của các đơn vị sản xuất do sử dụng những tài sản này trong quá trình sản xuất hoặc do hao mòn vô hình. Khấu hao tài sản cố định không bao gồm phần giá trị của tài sản bị mất đi do thiên tai, hỏa hoạn hoặc chiến tranh gây nên. Định nghĩa và mối liên hệ giữa thu nhập quốc gia thuần và GNI như sau:

4.2.12. Thu nhập quốc gia thuần

Thu nhập quốc gia thuần (NNI) phản ánh phần còn lại của tổng thu nhập quốc gia sau khi trừ đi khấu hao tài sản cố định dùng trong sản xuất của toàn bộ nền kinh tế trong một thời kỳ nhất định. Dưới dạng công thức, thu nhập quốc gia thuần được tính như sau:

$$\text{NNI} = \text{GNI} - \text{Khấu hao tài sản cố định dùng trong sản xuất}$$

Tổng thu nhập quốc gia hay thu nhập quốc gia thuần phản ánh thu nhập do sở hữu các yếu tố sản xuất của quốc gia tạo nên cho dù những yếu tố sản xuất này tham gia vào quá trình sản xuất ở trong nước hay bên ngoài. Trong quan hệ kinh tế và chính trị giữa các quốc gia, do tham gia vào hoạt động chung của cộng đồng quốc tế sẽ làm tăng hoặc giảm thu nhập của một quốc gia. Chẳng hạn, Việt Nam là thành viên của Liên Hợp Quốc, hàng năm phải nộp niên liễm sẽ làm giảm tổng thu nhập quốc gia của Việt Nam. Việt Nam viện trợ lương thực khẩn cấp cho nhân dân các nước bị sóng thần cũng làm giảm GNI của nước ta. Ngược lại, những người Việt nam sống xa tổ quốc hàng năm gửi kiều hối về cho gia đình và các thân nhân làm tăng tổng thu nhập quốc gia của Việt Nam, v.v... Tất cả các hoạt động giao dịch vừa nêu thuộc nhóm chuyển nhượng hiện hành.

Tổng thu nhập quốc gia hay thu nhập quốc gia thuần chưa phải là toàn bộ thu nhập của một nước dùng cho mục đích tiêu dùng cuối cùng hay để dành. Chỉ tiêu phản ánh thu nhập của quốc gia dùng cho

tiêu dùng cuối cùng và để dành của toàn bộ nền kinh tế là *thu nhập quốc gia khả dụng*. Định nghĩa và mối liên hệ giữa thu nhập quốc gia khả dụng và tổng thu nhập quốc gia như sau:

4.2.13. Thu nhập quốc gia khả dụng

Thu nhập quốc gia khả dụng phản ánh tổng thu nhập của quốc gia từ sản xuất, thu nhập sở hữu và từ chuyển nhượng hiện hành. Thu nhập quốc gia khả dụng dùng cho tiêu dùng cuối cùng và để dành và bằng thu nhập quốc gia cộng với chuyển nhượng hiện hành thuần từ bên ngoài.

Tên gọi thu nhập quốc gia khả dụng đã phản ánh thu nhập cuối cùng của nền kinh tế có được từ tất cả các nguồn để sử dụng cho hai mục đích: Tiêu dùng cuối cùng của hộ gia đình, của Nhà nước và để dành. Tiêu dùng cuối cùng là chỉ tiêu kinh tế tổng hợp quan trọng phản ánh toàn bộ chi tiêu về hàng hóa và dịch vụ của hộ gia đình và Nhà nước để đáp ứng nhu cầu về vật chất và tinh thần của cá nhân dân cư và toàn xã hội. Trái với tiêu dùng cuối cùng, để dành là nguồn tài chính dùng cho đầu tư tái sản xuất mở rộng của nền kinh tế trong các kỳ tiếp theo. Để dành càng lớn chứng tỏ nền kinh tế có tiềm lực phát triển kinh tế bền vững và ít phụ thuộc vào nguồn tài chính từ bên ngoài. Một cách tổng quát, để dành được định nghĩa như sau:

4.2.14. Để dành

Để dành là phần thu nhập còn lại của thu nhập quốc gia khả dụng sau khi trừ đi phần thu nhập sử dụng cho tiêu dùng cuối cùng. Công thức thường áp dụng tính để dành như sau:

$$\text{Để dành (Sn)} = \text{Thu nhập quốc gia khả dụng (NDI)} - \text{Tiêu dùng cuối cùng}$$

Cũng như thu nhập quốc gia, chỉ tiêu thu nhập quốc gia khả dụng và để dành cũng được tính dưới dạng gộp và thuần và chúng khác nhau bởi khấu hao tài sản cố định.

Để dành là nguồn tài chính trong nước quan trọng cho đầu tư phát triển sản xuất. Bên cạnh để dành trong nước, các thực thể của nền kinh tế còn thực hiện chuyển nhượng tài sản với bên ngoài qua việc chuyển ra nước ngoài hay nhận từ nước ngoài. Như đã đề cập trong phần một số khái niệm cơ bản, kết quả của chuyển nhượng tài sản làm thay đổi của cải của nền kinh tế. Vì vậy, để dành cộng với chuyển nhượng tài sản thuần từ bên ngoài phản ánh *thay đổi của cải thuần* do để dành và chuyển nhượng tài sản tạo ra.

Ngoài hai yếu tố để dành và chuyển nhượng tài sản thuần từ bên ngoài dẫn tới thay đổi của cải thuần của nền kinh tế, còn một số nguyên nhân khác như: Thay đổi giá cả, thay đổi về khối lượng tài sản do thiên tai, hỏa hoạn, v.v... Tuy vậy, tài khoản vốn tài sản trong Hệ thống tài khoản quốc gia chỉ đề cập tới yếu tố thay đổi do để dành thuần và chuyển nhượng tài sản mang lại. Yếu tố phản ánh thay đổi giá trị của cải thuần do thay đổi giá cả và nguyên nhân khác sẽ phản ánh trong bảng tổng kết tài sản của quốc gia. Định nghĩa tổng quát của chỉ tiêu thay đổi của cải thuần do thay đổi để dành và chuyển nhượng tài sản như sau:

4.2.15. Thay đổi của cải thuần do thay đổi để dành và chuyển nhượng tài sản

Thay đổi của cải thuần do thay đổi để dành và chuyển nhượng tài sản phản ánh nguồn tài chính của nền kinh tế dùng cho mục đích đầu tư làm tăng tích lũy tài sản cố định, tài sản lưu động, tài sản quý hiếm và tài sản phi tài chính không do sản xuất tạo ra. Với định nghĩa này đã gợi cho các nhà thống kê hai cách tiếp cận để tính chỉ tiêu thay đổi của cải thuần do thay đổi để dành và chuyển nhượng tài sản. Cách thứ nhất tiếp cận từ các yếu tố bên ngoài dẫn tới thay đổi của cải thuần, cách thứ hai tiếp cận qua các yếu tố bên trong sử dụng để làm tăng của cải vật chất của nền kinh tế.

Nhu cầu mở rộng phát triển sản xuất của nền kinh tế trong mỗi thời kỳ sẽ quyết định quy mô và độ lớn của tích lũy tài sản. Vì vậy

thay đổi của cải thuần do thay đổi để dành và chuyển nhượng tài sản (nguồn tài chính) có thể bù đắp đủ cho tích lũy tài sản, có thể không đủ hoặc dư thừa. Chênh lệch giữa thay đổi của cải thuần do thay đổi để dành và chuyển nhượng tài sản với tổng tích lũy tài sản và tài sản phi tài chính không do sản xuất tạo ra phản ánh nguồn tài chính còn lại của nền kinh tế có khả năng cho bên ngoài vay và được gọi là *cho vay thuần* nếu chênh lệch là dương. Khi chênh lệch là âm, điều đó chứng tỏ nguồn tài chính trong nước không đủ cho tích lũy tài sản, nền kinh tế phải vay từ bên ngoài và được gọi là *đi vay thuần*.

Khi quốc gia A cho quốc gia B vay, quốc gia A đã xác lập (sở hữu) một trái quyền tài chính với quốc gia B. *"Trái quyền tài chính là tài sản xác lập cho người sở hữu nó (chủ nợ) quyền được nhận một khoản thanh toán hay một loạt những thanh toán từ một đơn vị khác (người nợ) trong những trường hợp cụ thể được quy định trong hợp đồng giữa họ với nhau"*⁽¹⁾. Tài sản tài chính là khái niệm rộng hơn trái quyền tài chính và được định nghĩa như sau: *"Tài sản tài chính là tài sản dưới dạng trái quyền tài chính, vàng tiền, quyền rút vốn đặc biệt do Quỹ tiền tệ quốc tế phân bổ, cổ phiếu công ty và các công cụ tài chính kinh doanh ngoài bảng"*⁽²⁾.

4.2.16. Cho vay thuần hay đi vay thuần

Trên góc độ tài chính, cho vay thuần hay đi vay thuần phản ánh tình trạng sở hữu tài sản tài chính của một quốc gia với bên ngoài và được tính bằng chênh lệch giữa nhận thuần về tích sản tài chính với phát sinh thuần về tiêu sản tài chính. Nhận thuần về tích sản tài chính của một quốc gia với bên ngoài là toàn bộ giá trị tài sản tài chính do bên ngoài phát hành được sở hữu bởi quốc gia đó. Ngược lại, phát sinh thuần về tiêu sản tài chính là toàn bộ giá trị tài sản tài chính do quốc gia đó phát hành được sở hữu bởi bên ngoài.

Các chỉ tiêu từ giá trị sản xuất, giá trị tăng thêm, v.v... đến cho

⁽¹⁾ Mục 10.4, Tài khoản quốc gia 1993

⁽²⁾ Mục 10.5, Tài khoản quốc gia 1993

vay thuần hay đi vay thuần là những chỉ tiêu kinh tế tổng hợp áp dụng khái niệm biến kỳ, nghĩa là chỉ tính cho một khoảng thời gian nhất định. Chỉ tiêu cho vay thuần hay đi vay thuần là chỉ tiêu cân đối trong tài khoản vốn tài sản và cũng là chỉ tiêu cân đối cuối cùng phản ánh quá trình từ sản xuất đến phân phối thu nhập và sử dụng nguồn tài chính cho tích lũy tài sản.

Để hoàn chỉnh bức tranh của toàn bộ nền kinh tế, bên cạnh các chỉ tiêu kinh tế tổng hợp phản ánh kết quả hoạt động của một thời kỳ chúng ta nghiên cứu một số chỉ tiêu phản ánh giá trị tài sản và của cải của quốc gia tại các thời điểm quan trọng trong năm. Những chỉ tiêu này xuất hiện trong bảng tổng kết tài sản của Hệ thống tài khoản quốc gia.

4.2.17. Bảng tổng kết tài sản

Bảng tổng kết tài sản là bức tranh mô tả những yếu tố làm thay đổi giá trị và khối lượng tài sản, phản ánh giá trị tích sản và tiêu sản tại các thời điểm đầu và cuối của kỳ hạch toán. Bảng tổng kết tài sản mô tả đồng nhất thức hạch toán cơ bản liên kết giá trị của từng loại tích sản tại thời điểm đầu kỳ với giá trị có tại thời điểm cuối kỳ qua các hoạt động giao dịch và những thay đổi khác về giá trị tài sản⁽¹⁾.

4.2.18. Cửa cải thuần

Cửa cải thuần là chỉ tiêu cân đối trong bảng tổng kết tài sản, phản ánh mức giàu, nghèo của một quốc gia. Cửa cải thuần được tính bằng chênh lệch giữa tích sản và tiêu sản tại một thời điểm cụ thể, thường vào đầu kỳ và cuối kỳ hạch toán. Để xác định giá trị của cải thuần tại một thời điểm cần phải tính giá trị của từng loại tích sản và tiêu sản theo giá thị trường tại thời điểm đó.

Theo thời gian, tổng giá trị của tích sản và tiêu sản luôn thay đổi do các nguyên nhân sau:

⁽¹⁾ Mục 2.55. Phương pháp biên soạn hệ thống tài khoản quốc gia ở Việt Nam, Nhà xuất bản Thống kê, Hà Nội - 2003.

- Thay đổi giá trị do các hoạt động giao dịch tài sản mang lại như: Mua thêm hoặc thanh lý các loại tài sản phi tài chính, do sử dụng tài sản vào sản xuất được thể hiện qua khấu hao tài sản, do chuyển nhượng tài sản;

- Thay đổi về khối lượng tài sản không do các hoạt động giao dịch trong nền kinh tế mang lại: Phát hiện thêm tài nguyên thiên nhiên mới, thay đổi giá trị tài sản do thiên tai, hỏa hoạn, chiến tranh, do thay đổi cách phân loại gây nên;

- Do biến động về giá cả của tài sản. Tài khoản quốc gia chia biến động về giá cả thành hai loại: Biến động giá của từng loại tài sản và biến động mức giá chung của toàn bộ nền kinh tế.

4.2.19. Cửa cải thuần đầu kỳ

Cửa cải thuần đầu kỳ phản ánh giá trị của cải của quốc gia tại thời điểm đầu kỳ hạch toán, được tính bằng chênh lệch giữa tích sản và tiêu sản đầu kỳ.

Mối liên hệ giữa tích sản /tiêu sản cuối kỳ với tích sản /tiêu sản đầu kỳ thể hiện qua các chỉ tiêu phản ánh ba nhóm nguyên nhân làm thay đổi giá trị của tích sản và tiêu sản vừa nêu ở trên. Cụ thể mối liên hệ của tích sản / tiêu sản giữa cuối kỳ và đầu kỳ như sau:

4.2.20. Tích sản phi tài chính cuối kỳ

Nhóm nguyên nhân thứ nhất làm thay đổi tích sản phi tài chính hình thành từ các hoạt động giao dịch làm tăng tích lũy tài sản trong khoảng thời gian từ đầu kỳ tới cuối kỳ, gồm tích lũy gộp tài sản cố định, tích lũy tài sản lưu động, tích lũy tài sản quý hiếm và chênh lệch tài sản phi tài chính không do sản xuất tạo ra. Tài sản vật chất hữu hình dùng vào sản xuất dẫn tới giá trị tài sản giảm đi và được thể hiện qua chỉ tiêu Khấu hao tài sản cố định. Vì vậy, giá trị tích sản phi tài chính cuối kỳ phải trừ đi giá trị khấu hao tài sản cố định trong kỳ. Nhóm nguyên nhân thứ hai làm thay đổi giá trị của tích sản phi tài

chính là do thay đổi khác về khối lượng tài sản và nhóm thứ ba được quy cho biến động về giá cả của tài sản vì vậy cần phải đánh giá lại giá trị của tài sản với mức giá phổ biến trên thị trường tại thời điểm cuối kỳ. Một cách tổng quát, công thức tính giá trị của tích sản phi tài chính cuối kỳ như sau:

Tích sản phi tài chính cuối kỳ bằng tích sản phi tài chính đầu kỳ cộng với tích lũy gộp tài sản cố định trừ đi khấu hao tài sản cố định cộng với Tích lũy tài sản lưu động cộng với tích lũy tài sản quý hiếm cộng với chênh lệch tài sản phi tài chính không do sản xuất tạo ra cộng với thay đổi khác về khối lượng tài sản phi tài chính cộng với đánh giá lại tài sản phi tài chính.

4.2.21. Tích sản tài chính cuối kỳ

Giống như tích sản phi tài chính cuối kỳ, có ba nguyên nhân dẫn tới thay đổi giá trị của tích sản tài chính đầu kỳ đó là nhận thuần về tài sản tài chính như sở hữu thêm hay giảm đi các tài sản tài chính do bên ngoài phát hành. Nguyên nhân thứ hai biểu thị những thay đổi khác về khối lượng tài sản tài chính như được phân bổ thêm hay bị xoá bỏ quyền rút vốn đặc biệt (SDR) từ Quỹ tiền tệ quốc tế. Nguyên nhân cuối cùng thuộc về đánh giá lại giá trị các tài sản tài chính do giá cả của chúng thay đổi trên thị trường vì những lý do kinh tế và chính trị. Một cách tổng quát, công thức tính giá trị của tích sản tài chính cuối kỳ như sau:

Tích sản tài chính cuối kỳ bằng tích sản tài chính đầu kỳ cộng với thay đổi khác về khối lượng tài sản tài chính cộng với đánh giá lại tài sản tài chính.

4.2.22. Tiêu sản cuối kỳ

Tương tự như tích sản cuối kỳ, có ba nhóm nguyên nhân biểu thị sự khác biệt giữa tiêu sản cuối kỳ và tiêu sản đầu kỳ. Công thức tính giá trị của tiêu sản cuối kỳ như sau:

Tiêu sản cuối kỳ bằng tiêu sản đầu kỳ cộng với tiêu sản thuần trong kỳ cộng với thay đổi khác về khối lượng của tiêu sản cộng với đánh giá lại tiêu sản.

4.2.23. Cửa cải thuần cuối kỳ

Cửa cải thuần cuối kỳ phản ánh giá trị của cải của một quốc gia tại thời điểm cuối kỳ hạch toán. Qua mối liên hệ giữa cửa cải thuần với tích sản và tiêu sản, cửa cải thuần cuối kỳ có thể tính bằng hai cách:

Cửa cải thuần cuối kỳ bằng chênh lệch giữa tích sản cuối kỳ và tiêu sản cuối kỳ và cũng bằng cửa cải thuần đầu kỳ cộng với thay đổi của cải thuần trong kỳ.

4.2.24. Thay đổi của cải thuần

Thay đổi của cải thuần là chỉ tiêu cuối cùng trong bảng tổng kết tài sản được đề cập ở đây và bằng tổng của ba loại thay đổi: Thay đổi của cải thuần do để dành và chuyển nhượng vốn, thay đổi của cải thuần do thay đổi khác về khối lượng tài sản và thay đổi của cải thuần do chênh lệch giá danh nghĩa. Ba loại thay đổi này được phân loại ứng với ba nguyên nhân dẫn tới thay đổi giá trị của tích sản và tiêu sản của nền kinh tế.

Trên đây đã trình bày nội dung và mối liên hệ giữa các chỉ tiêu kinh tế tổng hợp của toàn bộ nền kinh tế từ khâu sản xuất tạo ra thu nhập đến khâu phân phối thu nhập, sử dụng thu nhập cho tiêu dùng, để dành, tích lũy và những thay đổi dẫn tới biến động của tích sản, tiêu sản cũng như của cải thuần tại thời điểm cuối kỳ của nền kinh tế. Toàn bộ mối liên hệ trên được mô tả dưới dạng phương trình liên hệ như sau:

Phương trình liên hệ của một số chỉ tiêu kinh tế tổng hợp

1. Tổng sản phẩm trong nước (GDP) theo giá thị trường

= Giá trị sản xuất theo giá cơ bản + Thuế, trừ trợ cấp sản phẩm + Thuế nhập khẩu hàng hóa và dịch vụ – Chi phí trung gian.

2. GDP theo giá thị trường

= Tiêu dùng cuối cùng + Tích lũy tài sản cố định + Tích lũy tài sản lưu động + Tích lũy tài sản quý hiếm + Xuất khẩu hàng hóa và dịch vụ – Nhập khẩu hàng hóa và dịch vụ.

3. GDP theo giá thị trường

= Thu nhập của người lao động từ sản xuất + Thuế sản xuất + Khấu hao tài sản cố định dùng trong sản xuất + Thặng dư

4. Tổng thu nhập quốc gia (GNI) theo giá thị trường

= GDP theo giá thị trường + Thuế, trừ trợ cấp sản xuất và hàng nhập khẩu thuần với bên ngoài + Thu nhập thuần của người lao động từ bên ngoài + Thu nhập sở hữu thuần với bên ngoài.

5. Thu nhập quốc gia thuần (NNI) theo giá thị trường

= GNI theo giá thị trường – Khấu hao tài sản cố định.

6. Thu nhập quốc gia khả dụng thuần (NNDI)

= Tiêu dùng cuối cùng + Để dành thuần.

7. Thay đổi của cải thuần do để dành và chuyển nhượng tài sản

= Để dành thuần + Chuyển nhượng tài sản thuần từ bên ngoài.

8. Để dành thuần + Chuyển nhượng tài sản thuần từ bên ngoài

= Tích lũy tài sản cố định – Khấu hao tài sản cố định + Tích lũy tài sản lưu động + Tích lũy tài sản quý hiếm + Chênh lệch tài sản phi tài chính không do sản xuất tạo ra + Cho vay thuần /- Đi vay thuần.

9. Cho vay thuần (+)/Đi vay thuần § (-)

= Nhận thuần về tích sản tài chính – Phát sinh thuần về tiêu sản tài chính.

10. Cửa cải thuần đầu kỳ

= Tích sản đầu kỳ – Tiêu sản đầu kỳ.

11. **Tích sản phi tài chính cuối kỳ**

= Tích sản phi tài chính đầu kỳ + Tích lũy gộp tài sản cố định – Khấu hao tài sản cố định + Tích lũy tài sản lưu động + Tích lũy tài sản quý hiếm + Chênh lệch tài sản phi tài chính không do sản xuất tạo ra + Thay đổi khác về khối lượng tài sản phi tài chính + Đánh giá lại tài sản phi tài chính.

12. **Tích sản tài chính cuối kỳ**

= Tích sản tài chính đầu kỳ + Nhận thuần tài sản tài chính + Thay đổi khác về khối lượng tài sản tài chính + Đánh giá lại tài sản tài chính.

13. **Tiêu sản cuối kỳ**

= Tiêu sản đầu kỳ + Tiêu sản thuần trong kỳ + Thay đổi khác về khối lượng của tiêu sản + Đánh giá lại tiêu sản.

14. **Của cải thuần cuối kỳ**

= Tích sản cuối kỳ – Tiêu sản cuối kỳ
= Của cải thuần đầu kỳ + Thay đổi của cải thuần trong kỳ

15. **Thay đổi của cải thuần**

= Thay đổi của cải thuần do để dành và chuyển nhượng vốn + Thay đổi của cải thuần do thay đổi khác về khối lượng tài sản + Thay đổi của cải thuần do chênh lệch giá danh nghĩa.

PHẦN NĂM

**MỘT SỐ CHỈ TIÊU THỐNG KÊ
KINH TẾ - XÃ HỘI TỔNG HỢP**

Ngoài các chỉ tiêu chủ yếu trong hệ thống tài khoản quốc gia, còn nhiều chỉ tiêu thống kê phản ánh khái quát, liên kết nhiều mặt kinh tế - xã hội hoặc phản ánh tập trung những mặt bản chất cho phép đánh giá sâu sắc mối quan hệ bên trong và tính quy luật của hiện tượng (gọi chung là các chỉ tiêu thống kê kinh tế - xã hội).

Trong số các chỉ tiêu đó, ở phần này của cuốn sách chỉ giới thiệu một số chỉ tiêu hiện nay đang được nhiều người quan tâm và rất cần thiết cho nghiên cứu, ứng dụng thực tế trong công tác thống kê.

5.1. HỆ SỐ ICOR

Hiệu quả sử dụng vốn đầu tư (viết tắt là ICOR) là chỉ tiêu kinh tế tổng hợp cho biết để tăng thêm một đơn vị tổng sản phẩm trong nước đòi hỏi phải tăng thêm bao nhiêu đơn vị vốn đầu tư thực hiện. Vì vậy, hệ số này phản ánh hiệu quả của việc sử dụng vốn đầu tư dẫn tới tăng trưởng kinh tế. Vốn đầu tư thực hiện trong hệ số ICOR bao gồm các khoản chi tiêu để làm tăng tài sản cố định, tài sản lưu động và các khoản hình thành nên giá trị tăng thêm của các ngành kinh tế. Hệ số ICOR thay đổi tùy theo thực trạng kinh tế - xã hội trong từng thời kỳ khác nhau, phụ thuộc vào cơ cấu đầu tư và hiệu quả sử dụng các sản phẩm vật chất và dịch vụ trong nền kinh tế.

Hệ số ICOR thấp chứng tỏ đầu tư có hiệu quả cao, hệ số ICOR thấp hơn có nghĩa là để duy trì cùng một tốc độ tăng trưởng kinh tế cần một tỉ lệ vốn đầu tư so với tổng sản phẩm trong nước thấp hơn. Tuy nhiên, theo quy luật về lợi tức biên giảm dần khi nền kinh tế

càng phát triển (GDP bình quân đầu người tăng lên) thì hệ số ICOR sẽ tăng lên, tức là để duy trì cùng một tốc độ tăng trưởng cần một tỉ lệ vốn đầu tư so với tổng sản phẩm trong nước cao hơn.

Có hai phương pháp tính hệ số ICOR:

* Phương pháp thứ nhất được tính theo công thức:

$$ICOR = \frac{V_1}{G_1 - G_0} \quad ; \quad (5.1.1)$$

Trong đó:

V_1 : Tổng vốn đầu tư của năm báo cáo;

G_1 : Tổng sản phẩm trong nước của năm báo cáo;

G_0 : Tổng sản phẩm trong nước của năm trước năm báo cáo.

Các chỉ tiêu về vốn đầu tư và tổng sản phẩm trong nước để tính hệ số ICOR theo phương pháp này phải được tính theo cùng một loại giá: Giá thực tế hoặc giá so sánh. Phương pháp tính thể hiện: Để tăng thêm đơn vị tổng sản phẩm trong nước, đòi hỏi phải tăng thêm bao nhiêu đơn vị vốn đầu tư thực hiện.

* Phương pháp thứ hai được tính theo công thức:

$$ICOR = \frac{I_V(\%)}{I_G(\%)} \quad ; \quad (5.1.2)$$

Trong đó:

I_V : Tỷ lệ vốn đầu tư so với tổng sản phẩm trong nước;

I_G : Tốc độ tăng tổng sản phẩm trong nước;

Hệ số ICOR tính theo phương pháp này thể hiện: Để tăng thêm 1 phần trăm (%) tổng sản phẩm trong nước đòi hỏi phải tăng bao nhiêu phần trăm tỷ lệ vốn đầu tư so với GDP.

Ví dụ: Năm 1996 Việt Nam có mức tăng tuyệt đối về GDP ($G_1 - G_0$) theo giá 1994 là 18266 tỷ đồng, tốc độ tăng GDP (I_G) là 9,34%;

vốn đầu tư phát triển theo giá so sánh (V_1) là 74134 tỷ đồng; tỷ trọng vốn đầu tư phát triển so với GDP (I_V) là 0,3212.

Theo số liệu trên:

- Áp dụng công thức 5.1.1 tính được:

$$ICOR = 74134 : 18266 = 4,068$$

- Áp dụng công thức 5.1.2 tính được:

$$ICOR = 0,3212 : 0,0934 = 3,439$$

Hai phương pháp tính hệ số ICOR nêu trên cho kết quả không giống nhau (kết quả tính toán ở trên có hệ số ICOR là 4,068 và 3,439). Trong thực tế người ta thường sử dụng phương pháp thứ nhất tính theo giá so sánh vì phương pháp này hạn chế được sai số thống kê và loại trừ ảnh hưởng của yếu tố giá một cách tốt nhất.

5.2. CHỈ SỐ PHÁT TRIỂN CON NGƯỜI

Chỉ số phát triển con người (HDI) là thước đo tổng hợp phản ánh sự phát triển của con người trên các phương diện thu nhập (thể hiện qua tổng sản phẩm trong nước bình quân đầu người), tri thức (thể hiện qua chỉ số học vấn) và sức khỏe (thể hiện qua tuổi thọ bình quân tính từ lúc sinh). Chỉ số phát triển con người được tính theo công thức:

$$HDI = \frac{1}{3}(HDI_1 + HDI_2 + HDI_3) \quad ; \quad (5.2)$$

Trong đó:

HDI_1 - Chỉ số GDP bình quân đầu người (GDP tính theo phương pháp sức mua tương đương "PPP" có đơn vị tính là đô la Mỹ);

HDI_2 - Chỉ số học vấn được tính bằng cách bình quân hóa giữa chỉ số tỷ lệ biết chữ (dân cư biết đọc, biết viết) với quyền số là 2/3 và chỉ số tỷ lệ người lớn (24 tuổi trở lên) đi học với quyền số là 1/3;

HDI_3 - Chỉ số tuổi thọ bình quân tính từ lúc sinh (kỳ vọng sống tính từ lúc sinh).

HDI nhận giá trị từ 0 đến 1. HDI càng gần 1 có nghĩa là trình độ phát triển con người càng cao, trái lại càng gần 0 nghĩa là trình độ phát triển con người càng thấp.

Công thức tính các chỉ số thành phần (HDI₁, HDI₂, HDI₃) như sau:

$$HDI_1 = \frac{\lg(\text{GDP thực tế}) - \lg(\text{GDP min})}{\lg(\text{GDP max}) - \lg(\text{GDP min})}$$

Từng chỉ số về tỷ lệ biết chữ và tỷ lệ người lớn đi học được tính toán riêng biệt nhưng đều theo công thức khái quát sau đây:

$$HDI_2 = \frac{L \text{ thực tế} - L \text{ min}}{L_{\text{max}} - L \text{ min}}$$

Ở đây: L - tỷ lệ người lớn đi học hoặc tỷ lệ biết chữ của dân cư.

$$HDI_3 = \frac{T \text{ thực tế} - T \text{ min}}{T_{\text{max}} - T \text{ min}}$$

Ở đây: T - Tuổi thọ bình quân tính từ lúc sinh

Các giá trị tối đa (max) và tối thiểu (min) của các chỉ tiêu liên quan để tính HDI được quy định như sau:

Chỉ tiêu	Đơn vị tính	Giá trị tối đa (max)	Giá trị tối thiểu (min)
GDP thực tế bình quân đầu người (PPP)	USD	40 000	100
Tỷ lệ dân cư biết chữ	%	100	0
Tỷ lệ người lớn đi học	%	100	0
Tuổi thọ bình quân tính từ lúc sinh	Năm	85	25

Ví dụ: Năm 1997 các chỉ tiêu chủ yếu của Việt Nam như sau:

- GDP thực tế bình quân đầu người (PPP) : 1 630 USD
- Tỷ lệ dân cư biết chữ : 91,9%

- Tỷ lệ người lớn đi học : 62,0%
- Tuổi thọ b /q tính từ lúc sinh : 67, 4 năm

Áp dụng công thức tính HDI nêu trên lần lượt tính các chỉ số thành phần qua số liệu đã cho như sau:

+ Chỉ số GDP bình quân đầu người:

$$HDI_1 = \frac{\lg(1630) - \lg(100)}{\lg(40000) - \lg(100)} = 0,466 = 0,466$$

+ Chỉ số học vấn (HDI₂):

$$- HDI_{2(b)} = \frac{91,9 - 0}{100 - 0} = 0,919 \text{ (chỉ số tỷ lệ biết chữ)}$$

$$- HDI_{2(d)} = \frac{62 - 0}{100 - 0} = 0,62 \text{ (chỉ số tỷ lệ đi học)}$$

$$- HDI_2 = \frac{1}{3} (0,62 + 2 \times 0,919) = 0,819 \text{ hoặc } 81,9\%$$

+ Chỉ số tuổi thọ:

$$HDI_3 = HD_3 \frac{67,4 - 25}{85 - 25} = 0,707$$

+ Chỉ số phát triển con người của Việt Nam vào năm 1997:

$$HDI = \frac{0,466 + 0,819 + 0,707}{3} = 0,664$$

5.3. CHỈ SỐ PHÁT TRIỂN GIỚI

Chỉ số phát triển giới (GDI) là thước đo phản ánh sự bất bình đẳng giữa nam và nữ trên cơ sở đánh giá sự phát triển chung của con người theo các yếu tố thu nhập, tri thức và tuổi thọ. Chỉ số phát triển giới được tính theo công thức:

$$GDI = \frac{1}{3}(GDI_1 + GDI_2 + GDI_3) \quad ; \quad (5.3)$$

Trong đó:

GDI_1 - Chỉ số phân bổ công bằng thành phần theo yếu tố thu nhập;

GDI_2 - Chỉ số phân bổ công bằng thành phần theo yếu tố tri thức;

GDI_3 - Chỉ số phân bổ công bằng thành phần theo yếu tố tuổi thọ.

GDI nhận giá trị từ 0 đến 1. GDI càng gần 1 có nghĩa là trình độ phát triển của con người trên góc độ bình đẳng về giới càng cao. Ngược lại càng gần 0 nghĩa là trình độ phát triển của con người trên góc độ bình đẳng về giới càng thấp.

Các chỉ số phân bổ công bằng thành phần theo các yếu tố thu nhập (1), tri thức (2) và tuổi thọ (3) viết chung là $GDI_{1(2,3)}$ được tính theo công thức:

$$GDI_{1(2,3)} = \left\{ \left[K^f (HDI_{1(2,3)}^f)^{1-\varepsilon} \right] + \left[K^m (HDI_{1(2,3)}^m)^{1-\varepsilon} \right] \right\}^{\frac{1}{1-\varepsilon}}$$

Trong đó:

f - Ký hiệu cho nữ và m - ký hiệu cho nam;

K^f - Tỷ lệ dân số nữ;

K^m - Tỷ lệ dân số nam.

$HDI_{1(2,3)}^f$ và $HDI_{1(2,3)}^m$ - Các chỉ số thành phần về từng yếu tố thu nhập (1), tri thức (2) và tuổi thọ (3) theo HDI của riêng nam và nữ.

ε - Hệ số phản ánh mức độ thiệt hại về phương diện phát triển con người mà xã hội gánh chịu do sự bất bình đẳng về giới. Trong chỉ số phát triển giới hệ số $\varepsilon = 2$ nên phương trình trên biến đổi thành:

$$GDI_{1(2,3)} = \left\{ \left[K^f (HDI_{1(2,3)}^f)^{-1} \right] + \left[K^m (HDI_{1(2,3)}^m)^{-1} \right] \right\}^{-1} \quad (*)$$

Tính chỉ số phát triển giới được thực hiện qua 3 bước:

Bước 1: Tính các chỉ số HDI thành phần riêng cho từng giới nữ và nam

Bước 2: Tính các chỉ số công bằng thành phần theo từng yếu tố thu nhập (GDI_1), tri thức (GDI_2) và tuổi thọ (GDI_3) theo công thức trên (*)

Bước 3: Tính chỉ số phát triển giới bằng cách bình quân số học giản đơn giữa 3 chỉ số phân bổ công bằng thành phần về thu nhập (GDI_1), tri thức (GDI_2) và tuổi thọ (GDI_3).

Giá trị tối đa (max) và tối thiểu (min) của các chỉ tiêu liên quan để tính GDP cho riêng từng giới quy định như sau:

Chỉ tiêu	Đơn vị tính	Giá trị tối đa (max)	Giá trị tối thiểu (min)
GDP thực tế bình quân đầu người (PPP)	USD	40 000	100
Tỷ lệ dân cư biết chữ	%	100	0
Tỷ lệ người lớn đi học	%	100	0
Tuổi thọ bình quân tính từ lúc sinh:			
- Nữ	năm	87,2	27,5
- Nam	năm	82,5	22,5

Ví dụ minh họa cho quá trình tính chỉ số phát triển giới với một số chỉ tiêu qua số liệu giả định như sau:

	Đơn vị tính	Nữ	Nam

- GDP thực tế bình quân đầu người	USD	1278	1982
- Tỷ lệ dân cư biết chữ	%	90,5	92,5
- Tỷ lệ người lớn đi học	%	61,5	62,5
- Tuổi thọ bình quân tính từ lúc sinh	Năm	71,2	63,8
- Tỷ lệ dân số theo giới	%	50,9	49,1

Từ số liệu đã cho, lần lượt tính toán:

Bước 1: Tính các chỉ số thành phần theo HDI của riêng từng giới

+ Chỉ số thu nhập:

$$HDI_1^f = \frac{\lg(1278) - \lg(100)}{\lg(40000) - \lg(100)} = \frac{3,106 - 2,000}{4,602 - 2,000} = 0,585$$

$$HDI_1^m = \frac{\lg(1982) - \lg(100)}{\lg(40000) - \lg(100)} = \frac{3,297 - 2,000}{4,602 - 2,000} = 0,638$$

+ Chỉ số tri thức:

- Chỉ số biết chữ

$$HDI_{2(b)}^f = \frac{90,5 - 0}{100 - 0} = 0,905; \quad HDI_{2(b)}^m = \frac{92,5 - 0}{100 - 0} = 0,925$$

- Chỉ số đi học

$$HDI_{2(d)}^f = \frac{61,5 - 0}{100 - 0} = 0,615; \quad HDI_{2(d)}^m = \frac{62,5 - 0}{100 - 0} = 0,625$$

- Chỉ số tri thức

$$HDI_2^f = \frac{2}{3} \cdot 0,905 + \frac{1}{3} \cdot 0,615 = 0,808;$$

$$HDI_2^m = \frac{2}{3} \cdot 0,925 + \frac{1}{3} \cdot 0,625 = 0,825$$

+ Chỉ số tuổi thọ:

$$HDI_3^f = \frac{71,2 - 27,5}{87,2 - 27,5} = 0,732; \quad HDI_3^m = \frac{63,8 - 22,5}{82,5 - 22,5} = 0,688$$

Bước 2: Tính các chỉ số phân bổ công bằng thành phần

+ Về thu nhập:

$$GDI_1 = \{ [0,509(0,585)^{-1}] + [0,491(0,638)^{-1}] \}^{-1} = 0,610$$

+ Về tri thức:

$$GDI_2 = \{ [0,509(0,808)^{-1}] + [0,491(0,825)^{-1}] \}^{-1} = 0,816$$

+ Về tuổi thọ:

$$GDI_3 = \{ [0,509(0,732)^{-1}] + [0,491(0,688)^{-1}] \}^{-1} = 0,709$$

Bước 3: Tính chỉ số phát triển giới

$$GDI = \frac{0,610 + 0,816 + 0,709}{3} = 0,711$$

So với chỉ số phát triển con người, nội dung và quy trình tính chỉ số phát triển giới không phức tạp hơn mấy. Song, thực tế áp dụng khó khăn hơn, vì tất cả các chỉ tiêu cần tính đều phải tính riêng theo từng giới. Hiện nay, thống kê Việt Nam chưa tách đầy đủ các chỉ số phân bổ theo giới, đặc biệt chỉ số phân bổ thành phần theo yếu tố thu nhập.

5.4. CHỈ SỐ BÌNH ĐẲNG VỀ GIỚI

Chỉ số bình đẳng về giới (GEM) là thước đo phản ánh sự bất bình đẳng giữa nam và nữ trong các lĩnh vực hoạt động chính trị, lãnh đạo quản lý, kỹ thuật, chuyên gia và thu nhập. Chỉ số bình đẳng về giới được tính theo công thức sau:

$$GEM = \frac{1}{3} (EDEP_1 + EDEP_2 + EDEP_3); \quad (5.4)$$

Trong đó:

EDEP₁ - Chỉ số phân bố công bằng thành phần theo số đại biểu nam, nữ trong Quốc hội;

EDEP₂ - Chỉ số phân bố công bằng thành phần theo vị trí lãnh đạo, quản lý, kỹ thuật, chuyên gia và kinh tế;

EDEP₃ - Chỉ số phân bố công bằng thành phần theo thu nhập.

GEM nhận giá trị từ 0 đến 1, GEM càng gần 1 nghĩa là sự bình đẳng về giới theo các lĩnh vực hoạt động chính trị, lãnh đạo quản lý, kỹ thuật, chuyên gia và thu nhập càng cao và ngược lại càng gần 0 thì sự bình đẳng theo giới về các tiêu thức trên càng thấp.

+ Chỉ số phân bố công bằng thành phần theo số đại biểu nam, nữ trong quốc hội (EDEP₁) được tính như sau:

$$EDEP_1 = \frac{\{ [k^f (I^f)^{-1}] + [k^m (I^m)^{-1}] \}^{-1}}{50} \quad (*)$$

Trong đó:

f - Ký hiệu cho nữ tính bằng số lần;

m - Ký hiệu cho nam tính bằng số lần;

k^f và k^m - Tỷ lệ dân số nữ và nam

I^f và I^m - Tỷ lệ nữ và tỷ lệ nam là đại biểu trong quốc hội (khác với k^f và k^m, trong công thức (*)) I^f và I^m được tính bằng phần trăm).

+ Chỉ số phân bố công bằng thành phần theo vị trí lãnh đạo, quản lý, kỹ thuật và chuyên gia (EDEP₂) được tính bằng cách bình quân số học giản đơn giữa 2 chỉ số EDEP tính riêng cho tỷ lệ nữ, nam theo lãnh đạo, quản lý và tỷ lệ nữ, nam theo vị trí kỹ thuật và chuyên gia. Từng chỉ số EDEP riêng biệt này được tính như công thức tính EDEP₁ nêu trên (*).

+ Chỉ số phân bố công bằng thành phần theo thu nhập (EDEP₃) được tính theo công thức:

$$EDEP_3 = \{ [k^f (H^f)^{-1}] + [k^m (H^m)^{-1}] \}^{-1} \quad (**)$$

Với H^f, H^m là các chỉ số thu nhập của nữ và nam, trong công thức (**) H^f và H^m được tính bằng số lần như k^f và k^m.

Về nguyên tắc, chỉ số thu nhập trong GEM tính toán tương tự như chỉ số thu nhập trong chỉ số phát triển giới. Song, điểm khác biệt là dựa trên giá trị không điều chỉnh và không lấy logarit. Công thức tính chỉ số thu nhập như sau:

$$H = \frac{\text{GDP thực tế} - \text{GDP min}}{\text{GDP max} - \text{GDP min}}$$

Mức thu nhập bình quân đầu người có giá trị tối đa (max) là 40 000 USD và giá trị tối thiểu (min) là 100 USD.

Quy trình tính chỉ số bình đẳng theo giới được thực hiện theo ba bước:

Bước 1: Xác định các tỷ lệ về giới theo đại biểu trong Quốc hội, theo vị trí lãnh đạo, quản lý và theo vị trí kỹ thuật và chuyên gia của nữ (I^f), nam (I^m) và các chỉ số thu nhập của nữ (H^f), nam (H^m),... để tính các chỉ số công bằng thành phần;

Bước 2: Tính các chỉ số phân bố công bằng thành phần EDEP₁, EDEP₂ và EDEP₃;

Bước 3: Tính chỉ số GEM bằng cách tính bình quân số học giản đơn giữa ba chỉ số phân bố công bằng thành phần về đại diện trong Quốc hội (EDEP₁), theo lãnh đạo quản lý, kỹ thuật và chuyên gia (EDEP₂) và theo thu nhập (EDEP₃).

Ví dụ minh họa cho quá trình tính chỉ số GEM với một số chỉ tiêu qua số liệu giả định như sau:

Chỉ tiêu	Nữ	Nam
1. Tỷ lệ đại biểu trong Quốc hội (%)	9,7	90,3

2. Tỷ lệ vị trí lãnh đạo và quản lý (%)	24,3	75,7
3. Tỷ lệ vị trí kỹ thuật và chuyên gia (%)	42,4	57,6
4. GDP bình quân đầu người (USD)	2556	3964
5. Tỷ lệ dân số theo giới (%)	50,9	49,1

Từ số liệu đã cho lần lượt tính toán:

Bước 1: Chỉ số thu nhập

$$H^f = \frac{2556 - 100}{40000 - 100} = 0,0616; \quad H^m = \frac{3964 - 100}{40000 - 100} = 0,0968$$

Bước 2: Tính các chỉ số công bằng thành phần

+ Chỉ số công bằng thành phần theo đại biểu trong Quốc hội (EDEP₁)

$$EDEP_1 = \frac{\{[0,509(9,7)^{-1}] + [0,491(90,3)^{-1}]\}^{-1}}{50} = 0,3454$$

+ Chỉ số công bằng thành phần theo lãnh đạo, quản lý, kỹ thuật và chuyên gia (EDEP₂)

- Theo vị trí lãnh đạo và quản lý (EDEP₂¹)

$$EDEP_2^1 = \frac{\{[0,509(2,43)^{-1}] + [0,491(75,7)^{-1}]\}^{-1}}{50} = 0,7291$$

- Theo vị trí kỹ thuật và chuyên gia (EDEP₂²)

$$EDEP_2^2 = \frac{\{[0,509(42,4)^{-1}] + [0,491(57,6)^{-1}]\}^{-1}}{50} = 0,9742$$

- Theo lãnh đạo, quản lý, kỹ thuật và chuyên gia nói chung

$$EDEP_2 = \frac{1}{2} [0,7291 + 0,9742] = 0,85165$$

+ Chỉ số phân bổ công bằng thành phần theo thu nhập (EDEP₃)

$$EDEP_1 = \{[0,509(0,06155)^{-1}] + [0,491(0,0968)^{-1}]\}^{-1} = 0,07497$$

Bước 3: Tính chỉ số bình đẳng theo giới:

$$GEM = \frac{1}{3} (0,3454 + 0,85165 + 0,07497) = 0,424$$

Giống như tính chỉ số phát triển giới, khó khăn lớn nhất để tính chỉ số bình đẳng về giới ở Việt Nam là việc tách chỉ tiêu tổng sản phẩm trong nước bình quân đầu người theo giới. Ngoài ra, nguồn số liệu để tính các tỷ lệ tham gia lãnh đạo, quản lý, chuyên gia, kỹ thuật và nghiệp vụ của nữ và nam cũng chưa thật đầy đủ.

5.5. TỐC ĐỘ TĂNG NĂNG SUẤT CÁC NHÂN TỔ TỔNG HỢP

Năng suất các nhân tố tổng hợp (TFP) suy cho cùng là kết quả sản xuất mang lại do nâng cao hiệu quả sử dụng vốn và lao động (các nhân tố hữu hình), nhờ vào tác động của các nhân tố vô hình như đổi mới công nghệ, hợp lý hoá sản xuất, cải tiến quản lý, nâng cao trình độ lao động của công nhân, v.v... (gọi chung là các nhân tố tổng hợp).

Tốc độ tăng năng suất các nhân tố tổng hợp là tỷ lệ tăng lên của kết quả sản xuất do nâng cao năng suất tổng hợp chung (năng suất tính chung cho cả vốn và lao động). Đây là chỉ tiêu phản ánh đích thực và khái quát nhất hiệu quả sử dụng nguồn lực sản xuất, làm căn cứ quan trọng để đánh giá chất lượng tăng trưởng cũng như tính chất phát triển bền vững của kinh tế, là cơ sở để phân tích hiệu quả sản xuất xã hội, đánh giá tiến bộ khoa học công nghệ, đánh giá trình độ tổ chức, quản lý sản xuất,... của mỗi ngành, mỗi địa phương hay mỗi quốc gia.

Tốc độ tăng năng suất các nhân tố tổng hợp ($\dot{F}_{TFP}$) được tính theo công thức:

$$\dot{F}_{TFP} = \dot{F}_Y - (\alpha \cdot \dot{F}_K + \beta \cdot \dot{F}_L) \quad ; \quad (5.5.1)$$

Trong đó:

$\dot{F}_Y$ - Tốc độ tăng kết quả sản xuất (kết quả sản xuất là giá trị tăng

thêm đối với từng ngành kinh tế, từng đơn vị hoặc từng khu vực, từng địa phương là tổng sản phẩm trong nước đối với toàn bộ nền kinh tế quốc dân).

$\dot{K}$ - Tốc độ tăng vốn hoặc tài sản cố định;

$\dot{L}$ - Tốc độ tăng lao động làm việc;

α , β là hệ số đóng góp của vốn hoặc tài sản cố định và hệ số đóng góp của lao động ($\alpha + \beta = 1$).

Để áp dụng được công thức trên ta phải có số liệu về 3 chỉ tiêu:

- Giá trị tăng thêm đối với từng ngành, từng đơn vị hoặc từng khu vực, từng địa phương và tổng sản phẩm trong nước đối với toàn bộ nền kinh tế quốc dân tính theo giá so sánh (giá cố định);

- Vốn hoặc tài sản cố định tính theo giá so sánh (giá cố định);

- Lao động làm việc.

Số liệu về giá trị tăng thêm hoặc tổng sản phẩm trong nước và số liệu về lao động có thể sử dụng trực tiếp hoặc khai thác để tính toán từ số liệu có trong các cuốn niên giám thống kê hàng năm hoặc hệ thống số liệu tổng hợp của ngành Thống kê. Nhưng riêng chỉ tiêu vốn hoặc giá trị tài sản cố định thì phải có kế hoạch theo dõi, cập nhật và áp dụng phương pháp tính toán, xử lý thích hợp với từng ngành kinh tế, từng phạm vi tổng hợp khác nhau.

Các hệ số đóng góp của vốn hoặc tài sản cố định (α) và của lao động (β) có thể xác định được bằng phương pháp hạch toán hoặc bằng hàm sản xuất Cobb-Douglass.

5.5.1. Tính các hệ số α và β theo phương pháp hạch toán

Công thức tính hệ số β theo phương pháp hạch toán có dạng:

$$\beta = \frac{\text{Thu nhập đầy đủ của người lao động theo giá hiện hành}}{\text{Giá trị tăng thêm hoặc GDP theo giá hiện hành}} \quad (5.5.2)$$

Như vậy khi tính hệ số β theo phương pháp hạch toán thì ngoài 3 chỉ tiêu trên (giá trị tăng thêm hoặc tổng sản phẩm trong nước), vốn hoặc tài sản cố định và lao động làm việc còn phải có thêm số liệu về thu nhập đầy đủ của người lao động và giá trị tăng thêm hoặc tổng sản phẩm trong nước theo giá hiện hành. Số liệu về hai chỉ tiêu để tính hệ số này có trong bảng I - O. Đối với những năm ta không lập bảng I - O thì phải dựa vào số liệu thu thập của người lao động có trong báo cáo thống kê của năm đó và hệ số chênh lệch về thu nhập của lao động theo bảng I - O và số liệu có trong các báo cáo thống kê của năm có bảng I - O để ước lượng thu nhập đầy đủ.

Khi có được hệ số β , ta dễ dàng xác định được hệ số α ($\alpha = 1 - \beta$ vì $\alpha + \beta = 1$).

5.5.2. Tính các hệ số α và β theo hàm sản xuất Cobb - Douglass

Hàm sản xuất Cobb - Douglass có dạng cơ bản:

$$\tilde{Y} = P.K^\alpha.L^\beta \quad ; \quad (5.5.3)$$

Trong đó:

$\tilde{Y}$ - Giá trị lý thuyết về tổng sản phẩm trong nước hoặc giá trị tăng thêm;

P - Năng suất bình quân chung;

K - Vốn hoặc giá trị tài sản cố định;

L - Lao động làm việc;

α - Hệ số đóng góp của vốn hoặc giá trị tài sản cố định;

β - Hệ số đóng góp của lao động, với $\alpha + \beta = 1$.

Tham số P và các hệ số α , β có thể tính được nhờ vào hệ phương

trình chuẩn tắc được xây dựng trên cơ sở phương pháp bình quân nhỏ nhất và quá trình tính toán các tham số đã được tiến hành như sau:

- Đưa hàm số $Y = P \cdot K^\alpha \cdot L^\beta$ về dạng tuyến tính bằng cách lấy logarit hai vế:

$$\begin{aligned}\ln Y &= \ln P + \alpha \ln K + \beta \ln L \\ &= \ln P + \alpha \ln K + (1 - \alpha) \ln L \\ &= \ln P + \alpha(\ln K - \ln L) + \ln L \Rightarrow \\ \ln Y - \ln L &= \ln P + \alpha(\ln K - \ln L)\end{aligned}$$

Hoặc:

$$\begin{aligned}\ln \frac{Y}{L} &= \ln P + \alpha \ln \frac{K}{L} \\ \ln P_L &= \ln P + \alpha \ln X\end{aligned}\quad ; \quad (5.5.4a)$$

Trong đó:

P_L là năng suất lao động $\left(P_L = \frac{Y}{L}\right)$

X là mức trang bị vốn cho lao động $\left(X = \frac{K}{L}\right)$

Nếu đặt: $\ln P_L = U;$ (a)

$\ln P = a_0;$ (b)

$\alpha = a_1$

$\ln X = z$

Thì phương trình 5.4a có dạng:

$$U = a_0 + a_1 z; \quad (5.5.4b)$$

- Lập và giải hệ phương trình chuẩn tắc để tìm a_0 và a_1 như sau:

$$\begin{cases} n \cdot a_0 + a_1 \cdot \sum z = \sum u \\ a_0 \sum z + a_1 \cdot \sum z^2 = \sum uz \end{cases} \quad ; \quad (5.5.4c)$$

Từ số liệu thực tế của nhiều năm về giá trị tăng thêm hoặc tổng

sản phẩm trong nước, vốn hoặc tài sản cố định, lao động làm việc ta tính các chỉ tiêu năng suất lao động P_L và mức trang bị vốn cho lao động (X), lấy logarit hai chỉ tiêu này rồi lập bảng tính toán và thay kết quả vào hệ phương trình 5.5.4c, giải ra ta được các tham số a_0 và a_1 .

- Khi có a_0 và a_1 (α) dễ dàng tìm được P và β vì $\ln P = a_0$ và $\beta = 1 - \alpha$.

Dưới đây ta xem ví dụ tính toán tốc độ tăng năng suất các nhân tố tổng hợp khi đã có số liệu về tốc độ tăng giá trị tăng thêm ($\dot{Y}$), tốc độ tăng giá trị tài sản ($\dot{K}$), tốc độ tăng lao động làm việc ($\dot{L}$) và các hệ số đóng góp của tài sản cố định (α) và lao động (β) của ngành công nghiệp Việt Nam bình quân năm giai đoạn 1991 c- 2003 như sau:

$$\dot{Y} = 13,33\% ; \quad \dot{K} = 14,08\% \text{ và } \dot{L} = 5,68\%$$

$$\alpha = 0,528 \text{ và } \beta = 0,472$$

Từ số liệu trên, áp dụng công thức 5.5.1 ta tính được:

$$\begin{aligned}\dot{Y}_{TFP} &= 13,33 - (0,528 \cdot 14,08 + 0,472 \cdot 5,68) \\ &= 13,33 - (7,44 + 2,68) = 3,21(\%) \end{aligned}$$

Tiếp tục ta tính toán tỷ phần đóng góp của tốc độ tăng TSCĐ, tốc độ tăng lao động và tốc độ tăng năng suất các nhân tố tổng hợp đối với tốc độ tăng giá trị tăng thêm.

- Tỷ phần đóng góp của tăng tài sản cố định:

$$(7,44 : 13,33) = 0,5577 \text{ hoặc } 55,77\%$$

- Tỷ phần đóng góp của tăng lao động:

$$(2,68 : 13,33) = 0,2012 \text{ hoặc } 20,12\%$$

- Tỷ phần đóng góp của tăng TFP:

$$(3,21 : 13,33) = 0,2410 \text{ hoặc } 24,10\%$$

Tỷ phần đóng góp của các nhân tố đối với tỷ lệ tăng lên của giá trị tăng thêm phản ánh vai trò của từng nhân tố. Khi tỷ phần đóng góp do tăng TFP càng cao thì hiệu quả kinh tế càng lớn và ngược lại.

5.6. HIỆU QUẢ QUÁ TRÌNH

Hiệu quả quá trình là mức năng suất được tính trên một phần chi phí chung còn lại sau khi trừ đi các khoản chi phí về nguyên vật liệu chính, vật liệu phụ, nhiên liệu,... tương đương như chi phí trung gian và phần chi phí còn lại này tạm gọi là chi phí chế biến.

Mức hiệu quả quá trình (P_E) được xác định như sau:

$$P_E = \frac{Y}{Z - IC} = \frac{Y}{Z - d.Z} \quad ; \quad (5.6)$$

Trong đó: Y là giá trị tăng thêm; hoặc tổng sản phẩm trong nước

Z là tổng chi phí sản xuất kinh doanh;

IC là chi phí trung gian.

d là tỷ lệ chi phí trung gian trong tổng chi phí

Ví dụ: Năm 1998, ngành công nghiệp "A" có giá trị sản xuất (GO) là 73693 triệu đồng, giá trị tăng thêm (Y) là 25454 triệu đồng, tổng chi phí sản xuất (Z) là 55406 triệu đồng, tỷ lệ giá trị trung gian trong tổng chi phí (d) là 0,7273

Theo công thức 5.6, ta tính được hiệu quả quá trình:

$$P_E = \frac{25454}{55406 - 0,7273.55406} = \frac{25454}{15140} = 1684 \text{ (đồng/1000 đồng)}$$

Chỉ tiêu hiệu quả quá trình nói lên một đơn vị chi phí chế biến tạo ra được bao nhiêu đơn vị giá trị tăng thêm hoặc đơn vị tổng sản phẩm trong nước. P_E càng lớn nghĩa là hiệu quả càng cao và ngược lại. Chỉ tiêu hiệu quả quá trình bảo đảm sự thống nhất về phạm vi so sánh giữa tử số và mẫu số (không phụ thuộc vào phần chi phí trung gian chiếm trong giá trị sản xuất cũng như trong tổng chi phí sản xuất lớn hay nhỏ). Sự biến động của chỉ tiêu hiệu quả quá trình phản ánh biến động tổng hợp của hiệu quả sử dụng các yếu tố chi phí về tài sản cố định và chi phí về lao động làm việc.

Có thể nói, đây là một trong những chỉ tiêu hiệu quả phản ánh một cách tập trung và khái quát về chất lượng sản xuất kinh doanh, về kết quả hoạt động tài chính của một doanh nghiệp, một ngành hay toàn bộ nền kinh tế quốc dân.

Song theo cơ cấu giá trị, chi phí để tính hiệu quả quá trình là một bộ phận của giá trị tăng thêm, tức là giữa phần chi phí đó và giá trị tăng thêm có một sự đồng nhất nhất định, nên sự thay đổi của quan hệ này (tăng, giảm hiệu quả quá trình) phù hợp với sự thay đổi cơ cấu của chỉ tiêu giá trị tăng thêm hoặc tổng sản phẩm trong nước và phụ thuộc nhiều vào nội dung, phương pháp và trình độ hạch toán của các chỉ tiêu đầu ra này.

Mặt khác, do tính theo chi phí nên chỉ tiêu hiệu quả quá trình chỉ mới thể hiện được hiệu quả sử dụng một bộ phận của nguồn lực đã trực tiếp sử dụng và được kết chuyển vào sản phẩm sản xuất ra, chứ chưa phản ánh đầy đủ hiệu quả của cả việc huy động lẫn sử dụng các yếu tố nguồn lực đã có vào quá trình sản xuất.

Chỉ tiêu hiệu quả quá trình có thể áp dụng được cho cả 3 cấp: doanh nghiệp, ngành và toàn nền kinh tế quốc dân. Song phù hợp và có điều kiện nhất vẫn là áp dụng cho cấp doanh nghiệp.

5.7. CHỈ SỐ THÀNH TỰU CÔNG NGHỆ

Chỉ số thành tựu công nghệ (TAI) là một chỉ tiêu thống kê tổng hợp phản ánh tổng quan về thành tích của mỗi nước trong việc sáng tạo, phổ biến công nghệ và xây dựng cơ sở các kỹ năng của con người. Chỉ số này đo các thành tựu đạt được trên 4 yếu tố sau:

- *Thứ nhất:* Sáng tạo công nghệ được đo bằng số bằng phát minh sáng chế cấp cho dân định cư tính trên đầu người và bằng số phí thu được về quyền sở hữu trí tuệ và bản quyền ra nước ngoài tính trên đầu người.

- *Thứ hai:* Phổ biến các thành tựu công nghệ gần đây, được đo bằng số máy chủ internet tính trên đầu người và tỷ lệ xuất khẩu hàng

công nghệ trung bình và hàng công nghệ cao.

- *Thứ ba*: Phổ biến các thành tựu công nghệ trước kia, được đo bằng số điện thoại (kể cả điện thoại cố định và điện thoại di động) trên đầu người và tiêu thụ điện năng tính trên đầu người.

- *Thứ tư*: Các kỹ năng của con người được đo bằng trung bình số năm đến trường của người lớn (15 tuổi trở lên) và tổng tỷ lệ nhập học đại học khoa học.

Tương ứng với 4 yếu tố trên ta có 4 chỉ số thành phần:

Chỉ số sáng tạo công nghệ - ký hiệu là I_1 , chỉ số phổ biến thành tựu công nghệ gần đây - ký hiệu là I_2 , chỉ số phổ biến thành tựu công nghệ trước kia - ký hiệu là I_3 và chỉ số kỹ năng của con người - ký hiệu là I_4 .

Chỉ số thành tựu công nghệ (TAI) là giá trị trung bình số học giản đơn của 4 chỉ số thành phần:

$$TAI = \frac{1}{4}(I_1 + I_2 + I_3 + I_4); \tag{5.7}$$

Mỗi chỉ số thành phần được tính từ 2 chỉ số riêng biệt khác nhau và chúng đều có giá trị quan sát tối thiểu và tối đa (đây là các giá trị quan sát giới hạn đồng thời được quy định thống nhất).

Khi tính các chỉ số riêng biệt thuộc chỉ số thành phần thứ nhất (I_1), chỉ số thành phần thứ hai (I_2) và chỉ số thành phần thứ tư (I_4) áp dụng công thức:

$$\text{Chỉ số riêng biệt} = \frac{\text{Giá trị thực tế} - \text{Giá trị quan sát tối thiểu}}{\text{Giá trị quan sát tối đa} - \text{Giá trị quan sát tối thiểu}}$$

Khi tính chỉ số riêng biệt thuộc thành phần thứ ba (I_3) áp dụng công thức:

$$\text{Chỉ số riêng} = \frac{\lg(\text{giá trị thực tế}) - \lg(\text{giá trị quan sát tối thiểu})}{\lg(\text{giá trị quan sát tối đa}) - \lg(\text{giá trị quan sát tối thiểu})}$$

$$\text{biệt} = \frac{\lg(\text{giá trị quan sát tối đa}) - \lg(\text{giá trị quan sát tối thiểu})}{\lg(\text{giá trị quan sát tối đa}) - \lg(\text{giá trị quan sát tối thiểu})}$$

Giá trị của mỗi chỉ số cá biệt có giá trị nằm giữa khoảng từ 0 đến 1.

Để minh họa cho nội dung và cách tính chỉ số thành tựu công nghệ, ta xét ví dụ về các kết quả đạt được của Niu -di-lân giai đoạn 1997 - 2000 như bảng 5.7.

Bảng 5.7: Các giá trị giới hạn quy định và giá trị thực tế đạt được giai đoạn 1997 - 2002

Chỉ tiêu	Giá trị giới hạn		Giá trị thực tế
	Tối đa	Tối thiểu	
Số bằng sáng chế cấp cho dân trong nước (trên 1 triệu dân)	994	0	100
Số thu quyền sở hữu trí tuệ và phí bản quyền (USD trên 1000 người)	272,6	0	13,0
Số máy chủ internet (trên 1000 người)	232,4	0	146,7
Tỷ lệ xuất khẩu hàng công nghệ cao và trung bình (%trong tổng xuất khẩu hàng hoá)	80,8	0	15,4
Số điện thoại (cố định và di động trên 1000 người)	901	1	720
Tiêu thụ điện năng trên đầu người (Kw giờ)	6969	22	6969
Trung bình số năm đến trường (từ 15 tuổi trở lên)	12,0	0,8	11,7

Tổng tỷ lệ nhập học đại học khoa học	27,4	0,1	13,1
--------------------------------------	------	-----	------

Từ số liệu của bảng ta tiến hành tính TAI theo hai bước:

Bước 1: Tính các chỉ số riêng biệt và các chỉ số thành phần

+ Tính chỉ số sáng tạo công nghệ:

$$\text{Chỉ số phát minh công nghệ} = \frac{100 - 0}{994 - 0} = 0,104$$

$$\text{Chỉ số phí sở hữu và bản quyền} = \frac{13,0 - 0}{272,6 - 0} = 0,048$$

$$\text{Chỉ số sáng tạo công nghệ} = \frac{0,104 + 0,048}{2} = 0,076$$

+ Tính chỉ số phổ biến các thành tựu công nghệ gần đây:

$$\text{Chỉ số máy chủ internet} = \frac{146,7 - 0}{232,4 - 0} = 0,631$$

$$\text{Chỉ số xuất khẩu hàng công nghệ cao và trung bình} = \frac{15,4 - 0}{80,8 - 0} = 0,191$$

$$\text{Chỉ số phổ biến các thành tựu công nghệ gần đây} = \frac{0,631 + 0,191}{2} = 0,411$$

+ Tính chỉ số phổ biến thành tựu công nghệ trước kia:

$$\text{Chỉ số điện thoại} = \frac{\log(720) - \log(1)}{\log(901) - \log(1)} = 0,967$$

$$\text{Chỉ số tiêu thụ điện năng} = \frac{\log(6969) - \log(22)}{\log(6969) - \log(22)} = 1$$

$$\text{Chỉ số phổ biến thành tựu công nghệ trước kia} = \frac{0,967 + 1}{2} = 0,984$$

+ Tính chỉ số kỹ năng con người:

$$\text{Chỉ số trung bình số năm đến trường} = \frac{11,7 - 0,8}{12 - 0,8} = 0,973$$

$$\text{Chỉ số tổng tỷ lệ nhập học đại học khoa học} = \frac{13,1 - 0,1}{27,4 - 0,1} = 0,476$$

$$\text{Chỉ số kỹ năng con người} = \frac{0,973 + 0,476}{2} = 0,725$$

Bước 2: Tính chỉ số thành tựu công nghệ trên cơ sở các chỉ số thành phần:

$$\text{TAI} = \frac{0,076 + 0,411 + 0,984 + 0,725}{4} = 0,549$$

TAI sẽ giúp các nhà hoạch định chính sách có cách nhìn nhận rõ hơn đối với tình trạng công nghệ hiện tại của đất nước, từ đó xác định chiến lược công nghệ trong thời đại thông tin. Hơn nữa, chỉ số này giúp các quốc gia xác định vị trí một cách tương đối so với các nước khác, để biết mình đang ở đâu.

5.8. CHỈ SỐ NGHÈO TỔNG HỢP

Chỉ số nghèo tổng hợp được xây dựng và áp dụng riêng cho 2 nhóm các nước có trình độ phát triển khác nhau: Nhóm các nước đang phát triển và nhóm các nước phát triển mà trực tiếp là các nước thuộc Tổ chức Hợp tác kinh tế và phát triển (viết tắt là OECD).

* Chỉ số nghèo tổng hợp cho các nước đang phát triển phản ánh sự khốn cùng trên 3 khía cạnh cơ bản về sự phát triển của con người:

- Không có khả năng đảm bảo một cuộc sống lâu dài và khỏe mạnh. Khía cạnh này được đo bằng xác suất không thọ quá 40 tuổi của con người (ký hiệu là P_1)

- Hạn chế về kiến thức (sự hiểu biết và khả năng giao tiếp của con người), khía cạnh này được đo bằng tỷ lệ người lớn mù chữ (ký hiệu

là P_2)

- Thiếu hụt về vật chất, không được tiếp cận tới những nhu cầu sinh hoạt tối thiểu như nước sạch, dịch vụ y tế, vệ sinh,... khía cạnh này được đo bằng tỷ lệ dân cư không được sử dụng nước sạch (ký hiệu là P_3') và tỷ lệ trẻ em dưới 5 tuổi thiếu cân, suy dinh dưỡng (ký hiệu là P_3''): $P_3' + P_3'' = P_3$

Phù hợp với các nội dung trên có công thức tính chỉ số nghèo tổng hợp cho các nước đang phát triển (HPI-1) như sau:

$$HPI-1 = [1/3 (P_1^\alpha + P_2^\alpha + P_3^\alpha)]^{1/3}; \quad (5.8.1)$$

Trong đó: $\alpha = 3$

Ví dụ có số liệu của một quốc gia đang phát triển như sau:

- Xác suất không thọ quá 40 tuổi: $P_1 = 11,9\%$

- Tỷ lệ người lớn mù chữ: $P_2 = 16,8\%$

- Tỷ lệ người dân không được tiếp cận với nước sạch:

$$P_3' = 21\%$$

- Tỷ lệ trẻ em dưới 5 tuổi thiếu cân $P_3'' = 6\%$

Từ số liệu trên ta tính:

+ Giá trị trung bình tỷ lệ người dân không được tiếp xúc nguồn nước sạch và tỷ lệ trẻ em dưới 5 tuổi thiếu cân.

$$P_3 = 1/2(21+6) = 13,5(\%)$$

+ Chỉ số nghèo tổng hợp

$$HPI-1 = \left[\frac{1}{3} (11,9^3 + 16,8^3 + 13,5^3) \right]^{1/3} = 14,4$$

* Chỉ số nghèo tổng hợp cho các nước OECD phản ánh sự khốn cùng trên 4 khía cạnh cơ bản về sự phát triển của con người:

- Không có khả năng đảm bảo cuộc sống lâu dài và mạnh khỏe

được đo bằng xác suất không thọ quá 60 tuổi của con người (ký hiệu là P_1)

- Hạn chế về kiến thức (sự hiểu biết và khả năng giao tiếp của con người). Khía cạnh này được đo bằng tỷ lệ người lớn mù chữ chức năng - ký hiệu là P_2)

- Mức độ thiếu thôn của dân cư, khía cạnh này được đo bằng tỷ lệ dân cư sống dưới mức chuẩn nghèo về thu nhập - 50% trung vị thu nhập của hộ gia đình (ký hiệu là P_3)

- Không có việc làm được đo bằng tỷ lệ thất nghiệp dài hạn - từ 12 tháng trở lên (ký hiệu là P_4)

Phù hợp với các nội dung trên có công thức tính chỉ số nghèo tổng hợp cho các nước OECD (HPI-2) như sau:

$$HPI-2 = [1/4 (P_1^\alpha + P_2^\alpha + P_3^\alpha + P_4^\alpha)]^{1/\alpha}; \quad (5.8.2)$$

Trong đó: $\alpha = 3$

Ví dụ có số liệu của một quốc gia thuộc tổ chức hợp tác kinh tế phát triển như sau:

- Xác suất thọ không quá 60 tuổi: $P_1 = 9,1\%$

- Tỷ lệ người lớn mù chữ chức năng: $P_2 = 17,0\%$

- Tỷ lệ dân cư dưới đường nghèo thu nhập: $P_3 = 2,1\%$

- Tỷ lệ thất nghiệp dài hạn: $P_4 = 14,3\%$

Từ số liệu trên, áp dụng công thức 5.8.2, ta tính được:

$$HPI-1 = \left[\frac{1}{4} (9,1^3 + 17,0^3 + 2,1^3 + 14,3^3) \right]^{1/3} = 12,9$$

TÀI LIỆU THAM KHẢO

1. Từ điển thống kê, Tổng cục Thống kê, Hà Nội - 1977
2. Niên giám Thống kê 1995, 2000, 2001, 2002, 2003, Tổng cục Thống kê xuất bản.
3. Giáo trình lý thuyết Thống kê, NXB. Giáo dục, Hà Nội - 1996
4. Phương pháp biên soạn Hệ thống tài khoản quốc gia ở Việt Nam, NXB. Thống kê, Hà Nội - 2003;
5. Phương pháp biên soạn Hệ thống tài khoản quốc gia ở Việt Nam, NXB. Thống kê, Hà Nội - 1998;
6. Giáo trình Thống kê kinh tế, NXB. Thống kê, Hà Nội - 2000;
7. Kết quả nghiên cứu đề tài cấp Tổng cục: Nghiên cứu xây dựng hệ thống từ chuẩn thống kê Việt Nam, do TSKH. Lê Văn Toàn làm chủ nhiệm;
8. Báo cáo phát triển con người 2001, NXB. Chính trị Quốc gia, Hà Nội - 2001;
9. Tổng cục Thống kê, Một số thuật ngữ thống kê thông dụng, NXB. Thống kê, Hà Nội - 2004;
10. Thông tin Khoa học Thống kê số 3/2001, 4/2001, 5/2002, 6/2002, 3/2005, Chuyên san Thống kê địa phương 2002, Chuyên san các chỉ số thống kê tổng hợp 2004;
11. Trung tâm nghiên cứu thông tin tư liệu dân số (Lê Văn Duyệt chủ biên) "Điều tra chọn mẫu - lý luận và ứng dụng trong công tác dân số và kế hoạch hoá gia đình", NXB. Thống kê, Hà Nội - 1998;
12. TS. Tăng Văn Khiên, Chỉ số khối lượng sản phẩm công

ngiệp, NXB. Thống kê, Hà Nội - 2001;

13. TS. Tăng Văn Khiên, Điều tra chọn mẫu và ứng dụng trong công tác thống kê, NXB. Thống kê, Hà Nội - 2003;

14. PGS. TS. Tăng Văn Khiên, Tốc độ tăng năng suất các nhân tố tổng hợp - phương pháp tính và ứng dụng, NXB. Thống kê, Hà Nội - 2005;

15. System of National Accounts 1993;

16. Australian System of national Accounts: Concept, sources and Methods, 2000;

17. John Sloman. Economics second Edition, Prentice Hall, Harvester Wheatsheaf;

18. Leslie Kish, Survey sampling, Publishing house John Wiley & Sons, INC 1995.